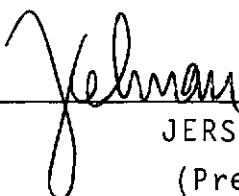


DETERMINAÇÃO DE VAZÕES EXTREMAS A PARTIR DE
SÉRIES PARCIAIS OU MÁXIMOS ANUAIS


Jonatas Costa Moreira

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE
PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JA-
NEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE MESTRE EM CIÊNCIAS (M.Sc.).

Aprovada por:



JERSON KELMAN
(Presidente)



PAULO CANEDO DE MAGALHÃES



BASÍLIO BRAGANÇA PEREIRA



BENEDITO PINTO FERREIRA BRAGA JR.

RIO DE JANEIRO, RJ-BRASIL

SETEMBRO DE 1983

MOREIRA, JONATAS COSTA

Determinação de vazões extremas a partir de séries parciais ou máximos anuais (Rio de Janeiro) 1983.

VIII, 152 p 29,7 cm (COPPE-UFRJ, M.Sc., Engenharia Civil, 1983)

Tese - Univ. Fed. Rio de Janeiro, COPPE.

I. Hidrologia, Vazões extremas, Séries Parciais
I. COPPE/UFRJ III. Título (série)

AGRADECIMENTOS

Ao Prof. Jerson Kelman pela sua orientação segura e apoio incondicional durante o desenvolvimento deste trabalho.

Ao CEPEL, à ELETROBRÁS e à CAPES, pelo apoio financeiro, indispensável à realização das pesquisas.

Aos colegas do Departamento de Sistemas do CEPEL, especialmente os pesquisadores Jorge Machado Damázio e Joarí Paulo da Costa, pelas sugestões, críticas e estímulos.

Ao Joarí e à Claudia, pela amizade e companheirismo com que compartilharam, não só o espaço de trabalho, como as alegrias e frustrações ao longo desta pesquisa.

Ao meu irmão José, pela elaboração das figuras e gráficos aqui apresentados.

Aos meus filhos Ulysses e Daniel, pelas inúmeras horas de lazer emprestadas.

Finalmente, a meus pais, Edgar e Rita, pelo apoio e carinho com que sempre acompanharam o desenvolvimento de meus estudos.

RESUMO

Esta tese tem como principal objetivo avaliar as potencialidades do uso de séries parciais para a análise de frequência de cheias. Em particular, os estudos são dirigidos para a estimação de vazões com tempos de recorrência muito elevados.

Inicialmente, comparam-se as principais propriedades das séries parciais com as das séries de máximos anuais correspondentes, tomando-se como base os registros de vazões médias diárias de 29 postos fluviométricos da bacia do rio Doce. A partir destas séries, são verificadas as principais hipóteses subjacentes aos modelos de séries parciais mais comuns.

É efetuada uma comparação direta entre os desempenhos dos métodos de séries parciais e os de máximos anuais na previsão de vazões com grandes tempos de recorrência. Para tanto, é utilizado um grande número de séries sintéticas de vazões médias diárias, geradas a partir de modelos estocásticos.

Finalmente, avalia-se a eficiência dos testes de adequação de ajuste, no tocante à escolha da distribuição de probabilidades mais apropriada para a estimação de vazões extremas.

ABSTRACT

The main purpose of this work is to evaluate the possibilities of the use of the Partial Duration Series (peaks over a threshold methods, POT) to estimate flood peaks with very high return periods.

Initially, the mean daily flow records of 29 gauging stations of the Rio Doce basin are used as a tool to verify the major assumptions of the POT methods. After that, the elementary properties of the Partial Series are compared to those of the correspondent Annual Maximum Series.

In order to compare the POT estimates of the T-year return period flood with the Annual Maximum estimates, synthetic hydrographs, obtained by stochastic models, are used.

That synthetic series are used as a basis to evaluate the efficiency of the goodness of fit index to selecting a probability distribution that best estimate floods with high return periods.

I N D I C E

	Página
CAPÍTULO I - INTRODUÇÃO	1
1.1 - Aspectos Gerais	1
1.2 - Análise de Frequência de Cheias	5
1.3 - Objetivos do Trabalho	10
CAPÍTULO II - ANÁLISE DIRETA DE FREQUÊNCIA DE CHEIAS	13
2.1 - Séries de Máximos Anuais e Séries Parciais..	13
2.2 - Modelos de Máximos Anuais	16
2.3 - Métodos de Séries Parciais	22
2.3.1 - Modelagem Matemática	22
2.3.2 - Critérios para Obtenção das Séries Parciais...	33
2.3.3 - A Aleatoriedade das Séries Parciais	37
CAPÍTULO III - SÉRIES DE VAZÕES EXTREMAS DA BACIA DO RIO DOCE	48
3.1 - Características Gerais das Séries de Extre- mos	48
3.1.1 - Estatísticas Locais e Regionais	49
3.1.2 - Modelos de Regressão Linear	54
3.2 - Hipóteses dos Modelos de Séries Parciais ...	57

3.2.1 - A Hipótese Poisson	57
3.2.2 - A Hipótese Exponencial	58
3.3 - Regionalização de Informações	65
3.3.1 - Distribuições Regionais de Frequência	66
3.3.2 - Homogeneidade da Bacia do Rio Doce	73
 CAPÍTULO IV - SELEÇÃO DE MÉTODOS PARA ESTIMAÇÃO DE VAZÕES COM PERÍODOS DE RECORRÊNCIA MUITO ELEVADOS..	 79
4.1 - Introdução	79
4.2 - Metodologia de Avaliação	84
4.3 - Aplicação a Séries Sintéticas: Modelo Xadrez ..	86
4.3.1 - Calibração do Modelo Xadrez	88
4.3.2 - Máximos Anuais x Séries Parciais	94
4.4 - Aplicação a Séries Sintéticas: Modelo Diana ..	110
4.5 - Testes de Adequação de Ajuste	118
 CAPÍTULO V - CONCLUSÕES E RECOMENDAÇÕES	 125
5.1 - Conclusões	125
5.2 - Recomendações	131
 BIBLIOGRAFIA	 134
 APÊNDICE A	 140
 APÊNDICE B	 146

CAPÍTULO I

INTRODUÇÃO

1.1 - Aspectos Gerais

A segurança de um aproveitamento hídrico está estreitamente ligada ao regime de vazões afluentes ao mesmo, uma vez que a ocorrência de enchentes excepcionais pode provocar danos que variam desde a simples inutilização de equipamentos, criando dificuldades à operação do sistema, até a ruptura total da obra, causando prejuízos às vezes irreparáveis. Ao projetar-se uma obra deste tipo é preciso fazer uma avaliação dos riscos de falha do sistema e dos investimentos necessários para evitá-los.

Com base nestas avaliações, podem ser determinadas medidas visando à redução dos riscos a valores admissíveis e adotadas certas precauções, como mecanismos de emergência, seguros contra cheias etc..

Uma medida do risco de falha de um aproveitamento hídrico pode ser dada pela probabilidade de ocorrência de uma vazão afluyente superior a um certo valor máximo admissível para o qual a obra foi projetada (vazão de projeto da obra), durante o período de utilização do aproveitamento. Para uma mesma vazão de projeto, o risco de falha do aproveitamento muda se forem considerados períodos de utilização diferentes, o que torna essa medida de risco pouco cômoda para o planejamento. Por isso, é prática comum em hidrologia tomar-se como medida de risco de falha de uma obra a probabilidade p de ocorrência, num ano qualquer, de uma cheia

cuja magnitude ultrapasse a vazão de projeto. O tempo de recorrência da vazão de projeto é dado pelo inverso dessa probabilidade p e corresponde ao valor esperado do intervalo entre a ocorrência de duas cheias sucessivas, com magnitudes superiores a esta vazão de projeto.

Existem obras cuja falha acarreta danos de ordem tão elevada que o risco admissível deve ser muito pequeno. Este é, por exemplo, o caso de uma grande barragem, cuja ruptura acarretaria prejuízos muito elevados, além de conseqüências não mensuráveis, como é o caso de perda de vidas humanas. A segurança de uma barragem está estreitamente ligada à capacidade de seus vertedores e ao regime de vazões afluentes, pois a ocorrência de vazões superiores à vazão de projeto dos vertedores pode acarretar o transbordamento da represa através de um ponto mais frágil da estrutura, ocasionando a sua ruptura completa. Como é praticamente impossível a construção de uma obra totalmente isenta de riscos, o que se faz neste caso é fixar riscos admissíveis muito pequenos, o que é equivalente à utilização de um vertedor cuja capacidade esteja associada a uma vazão com tempo de recorrência muito elevado.

A determinação de vazões com tempos de recorrência elevados pode ser feita através da análise de frequência de cheias, que procura associar uma lei probabilística adequada à variável aleatória "magnitude das vazões afluentes ao local de interesse". A análise de frequência de cheias pode considerar todo o tipo de informações hidrológicas, climatológicas e geomorfológicas existentes, sendo que a utilização de informações mais completas exige modelos matemáticos mais complexos e nem sempre conduz a resultados mais confiáveis.

Existem diversos modelos matemáticos para a análise de frequência de cheias, e a determinação de qual deles é o mais adequado a cada caso é sempre difícil, uma vez que a verdadeira população de vazões não é conhecida. Além disso, mesmo que se conhecesse o modelo correto a ser utilizado, a estimação de seus parâmetros poderia ser feita de formas diferentes, sendo a escolha da forma mais correta igualmente difícil.

Assim, a estimação de vazões com tempos de recorrência elevados podem estar associados dois tipos de erros:

- 1) Um erro decorrente do fato de o modelo a ser utilizado não representar fielmente a população em estudo, e
- 2) Um erro decorrente da estimação inadequada dos parâmetros do modelo a ser utilizado.

Inúmeras tentativas têm sido feitas no sentido de escolher modelos mais adequados e a melhor forma de se estimarem seus parâmetros; todas esbarram no pequeno tamanho das amostras (informações) disponíveis e no desconhecimento da população que lhes deu origem.

Chow (1964) considera que uma análise de frequências pode basear-se em dados experimentais ou em dados históricos. Quando se dispõe dos primeiros é possível chegar a resultados bastante conclusivos na análise, bastando para isso aumentar o número de experimentos, ou seja: aumentar o tamanho das amostras até que seja possível uma caracterização satisfatória da população. Quando se dispõe apenas de dados históricos, como é o caso da análise de frequência de cheias, as conclusões obtidas

são drasticamente limitadas pela extensão dos registros históricos de informações disponíveis.

Este trabalho procura, a partir da geração de inúmeras hidrógrafas sintéticas, fornecer dados "experimentais" que possam permitir a obtenção de resultados mais conclusivos, a serem aplicados na escolha de modelos para análise de frequência de cheias. São serão considerados os modelos que utilizam como informação apenas os registros de vazões afluentes ao local do aproveitamento.

Normalmente, são disponíveis dois tipos de registros de vazões:

- 1) Registros contínuos ao longo do tempo, obtidos a partir de limnógrafos, e
- 2) Registros discretos de vazões médias diárias.

Como no Brasil a disponibilidade de registros contínuos de vazões é muito pequena, são serão utilizadas aqui as séries de vazões médias diárias. Isso constitui uma limitação do trabalho, pois as vazões médias diárias são obtidas através de duas medições (geralmente, às 7:00hs e às 17:00hs), o que pode acarretar, para bacias com áreas de drenagem pequenas, a perda do pico de uma cheia e, até mesmo, de toda a cheia. Por isso, são foram consideradas bacias hidrográficas rurais com áreas de drenagem superiores a 100 km², sendo as conclusões obtidas aplicáveis apenas a estes casos.

1.2 - Análise de Frequência de Cheias

O objetivo da análise de frequência de cheias é obter uma relação entre uma magnitude de vazão q_T e o tempo de recorrência, T , a ela associado, para uma determinada seção de um curso d'água. O problema da análise de frequência de vazões extremas pode ser colocado da seguinte forma:

conhecidos os registros de dados hidrológicos de uma certa bacia hidrográfica, bem como um conjunto de características geomorfológicas e climatológicas da mesma, estabelecer uma lei que relacione as magnitudes das vazões num rio desta bacia com as suas probabilidades de ocorrência.

Existe um grande número de mecanismos físicos a cuja ocorrência estão associadas as vazões afluentes a uma determinada seção de um curso d'água que se quer estudar. Dentre eles, podem ser citados as características geomorfológicas da bacia hidrográfica e seu estado quando da ocorrência de chuvas, o regime de ocorrência de precipitações e sua distribuição espacial sobre a mesma, características meteorológicas agravantes etc. Os métodos que levam em conta todos esses mecanismos (ou vários deles), para a determinação de vazões associadas a diferentes tempos de recorrência (ou probabilidades), são chamados métodos de análise indireta de frequência de cheias. Estes métodos exigem, em geral, uma grande quantidade de dados, o que acarreta custos muito elevados ou, até mesmo, a impossibilidade de sua aplicação por carência de informações. Mais: como a cada um destes dados está associado um certo grau de incerteza, torna-se muito difícil a análise do desempenho de tais métodos.

Uma forma alternativa de tratar o problema da análise de frequência de cheias é considerar unicamente a série de vazões afluentes à seção em estudo como sendo uma realização de um processo estocástico cujos mecanismos de geração não são tratados um a um, mas, sim, de forma global. Nesta série são procuradas evidências estatísticas que possam caracterizar tal processo, cuja identificação permitiria resolver o problema. Os métodos que usam este procedimento são chamados métodos de análise direta de frequência de cheias. Constitui o objetivo deste estudo a investigação do comportamento de alguns destes métodos com relação à determinação de vazões com tempos de recorrência muito elevados, como é o caso das vazões de projeto de vertedores. Ocorre com frequência que os registros de vazões diárias disponíveis têm pequena extensão; o que acarreta grandes incertezas quanto ao resultado da análise. No Brasil, estes registros poucas vezes ultrapassam a casa dos 50 anos. Pode-se resumir a dificuldade na seguinte pergunta: dado um registro de vazões médias diárias com extensão de 30 anos, o que se pode afirmar acerca da vazão que é ultrapassada em média uma vez a cada 10.000 anos? Dada a pequena extensão do registro existente, face ao grande tempo de recorrência de interesse, a resposta a esta questão envolve um considerável grau de extrapolação, sendo por isso importante que dele se tente extrair a maior quantidade possível de informação.

Seria ideal que, na análise de frequência de cheias, fossem utilizadas todas as informações contidas no registro de vazões diárias. Todavia, esta série apresenta uma estrutura de dependência muito complexa, o que dificulta o tratamento analítico. Assim, o procedimento usualmente adotado na prática consiste em extrair-se destes registros de vazões diárias as chamadas séries

de vazões extremas, centrando, desta forma, a atenção nas vazões que apresentam interesse mais direto na análise de extremos.

Existem basicamente dois tipos de séries de vazões extremas: as séries de máximos anuais e as séries parciais. As primeiras são constituídas pelos maiores valores de vazões ocorridos a cada ano, enquanto que as últimas compreendem as maiores vazões ocorridas a cada cheia.

A simplicidade dos modelos matemáticos usados na a nálise de máximos anuais, aliada à facilidade com que estas séries são obtidas, faz com que, na prática, seu uso seja mais difundido que o de séries parciais, cuja análise exige o emprego de modelos matemáticos mais sofisticados. Além disso, a obtenção das séries parciais a partir dos registros de vazões diárias exige uma definição das cheias ocorridas no histórico, o que geralmente dá margem a controvérsias. Devido a esta dificuldade, costuma-se definir séries parciais como as séries contendo todos os picos da hidrografa cujas magnitudes excedam um determinado valor, usualmente chamado nível-base ou nível de truncamento.

Têm sido desenvolvidos estudos (NERC, 1975; Yevjevich e Taesombut, 1976) com o objetivo de comparar o desempenho dos métodos de análise de frequência de cheias a partir das séries de máximos anuais (método de máximos anuais) ou a partir das séries parciais (método das séries parciais). Uma aparente vantagem dos métodos de séries parciais é a possibilidade de se manipularem amostras maiores que nos métodos de máximos anuais, pois em um só ano pode ocorrer mais de uma cheia. Entretanto, existem dúvidas sobre a qualidade destas amostras, principalmente porque dois picos

de cheias consecutivos podem guardar entre si um elevado grau de dependência, o que anularia a aparente vantagem. Além disso, os métodos de máximos anuais consideram uma informação adicional: cada elemento da amostra é a maior vazão ocorrida em um ano (NERC, 1975). Os resultados destes estudos comparativos não têm sido conclusivos. Por um lado, as comparações feitas em bases analíticas são limitadas por inúmeras hipóteses simplificadoras, que acabam por comprometer as conclusões. Por outro lado, para uma comparação direta entre os métodos de máximos anuais e os de séries parciais seria preciso dispor de um número muito elevado de hidrôgrafas, das quais se pudesse obter os dois tipos de séries de extremos. Além disso, seria necessário o conhecimento da população que deu origem a estas séries.

A geração de um número elevado de hidrôgrafas sintéticas pode fornecer uma base objetiva para uma comparação direta entre o uso de séries parciais e de séries de máximos anuais, na análise de vazões extremas. Estas hidrôgrafas sintéticas constituiriam uma população perfeitamente conhecida (cenário), e os desempenhos dos dois métodos poderiam ser avaliados e comparados. As conclusões fornecidas por um estudo desta natureza podem ser aplicáveis, desde que os cenários criados apresentem características semelhantes à realidade que se tenta representar. Assim, se as séries de vazões diárias geradas não forem estatisticamente distinguíveis dos registros de vazões que lhes serviram de base, tal estudo comparativo pode fornecer subsídios para aplicação em casos reais.

Mesmo que se escolha o tipo de séries de extremos a ser utilizado na análise de frequência de cheias, é preciso decidir qual o modelo probabilístico mais adequado para aplicação em

cada caso. De maneira geral, a escolha do modelo a ser usado é fonte de inúmeras controvérsias, pois a aplicação de diferentes modelos pode conduzir a resultados muito distintos, principalmente se o estudo visar à determinação de vazões com tempos de recorrência muito elevados. Diversas metodologias têm sido propostas, no intuito de permitir maior confiabilidade na determinação destas vazões extremas, sendo freqüentes as propostas de uniformização de metodologias e a recomendação de distribuições de probabilidades padronizadas para aplicações regionais. Dois aspectos devem ser ressaltados aqui:

- 1) A grande maioria dos trabalhos sobre o assunto se aplica a séries de máximos anuais apenas, e
- 2) O critério de seleção da distribuição de probabilidades a ser usada é, geralmente, o de adequação de ajustes.

Considerando a pequena extensão dos registros disponíveis, um modelo probabilístico que se ajusta satisfatoriamente a uma amostra (ou seja, que reproduz com precisão a freqüência dos eventos verificados na mesma), pode não conduzir a extrapolações confiáveis, principalmente se este modelo não for adequado à população que forneceu aquela amostra.

Em vista destas considerações, seria interessante uma verificação experimental. O cenário de vazões diárias pode permitir também uma avaliação da eficiência dos testes de adequação de ajuste, como base para a escolha da distribuição de probabilidades mais adequada para a estimação de vazões com tempos de recorrência muito elevados, seja nos métodos de máximos anuais, seja nos de séries parciais.

1.3 - Objetivos do Trabalho

Tendo em vista os aspectos abordados nos dois itens anteriores, apresenta-se agora um apanhado da estrutura e dos objetivos deste trabalho. A estimação de vazões com tempos de recorrência muito elevados, tomando por base apenas os registros de vazões médias diárias, exige um cuidado especial para o aproveitamento da maior quantidade de informações que estes registros possam fornecer. Além disso, como a verdadeira distribuição de probabilidades da população não é conhecida, é forçosa a adoção de inúmeras hipóteses, a partir das quais possam ser feitas as extrapolações necessárias.

As séries de máximos anuais e as séries parciais são duas formas alternativas de se considerar as informações contidas num registro de vazões diárias. Entretanto, os métodos de séries parciais sô têm sido sugeridos para a estimação de vazões com tempos de recorrência muito pequenos. Ora, se não houvessem dúvidas quanto à escolha do modelo probabilístico correto, as séries parciais, por fornecerem amostras mais extensas, garantiriam mais confiança na determinação dos parâmetros do modelo e, conseqüentemente, na estimação de vazões extremas.

O principal objetivo deste trabalho é, então, investigar as potencialidades dos métodos de séries parciais para a análise de freqüência de cheias, particularmente na determinação de vazões com tempos de recorrência muito elevados.

O Capítulo II apresenta uma revisão dos modelos probabilísticos de máximos anuais e de séries parciais, com ênfase nestes últimos.

No Capítulo III, as hipóteses subjacentes a cada modelo são verificadas, através de alguns testes estatísticos aplicados às séries de vazões médias diárias de 29 postos da bacia do Rio Doce. Buscando caracterizar estes dois tipos de séries de extremos, são calculadas várias estatísticas locais e regionais. A possibilidade de utilização de informações regionais nos métodos das séries parciais e de máximos anuais é também discutida, através de alguns exemplos numéricos. Como estas informações regionais só têm sentido se a região em estudo apresentar características de homogeneidade do ponto de vista hidrológico, esta homogeneidade é testada através da verificação do ajuste de distribuições de probabilidades regionais aos vários postos, individualmente.

Uma comparação direta entre os métodos de séries parciais e os métodos de máximos anuais exige o conhecimento de uma população e a disponibilidade de amostras dos dois tipos de séries, obtidas a partir dessa mesma população. A geração de extensas séries sintéticas de vazões diárias pode permitir esta comparação. No Capítulo IV é apresentada uma metodologia para a avaliação da eficiência dos modelos de análise de frequência de cheias. O objetivo desta metodologia é responder às seguintes perguntas:

- 1) Qual o modelo que conduz aos menores erros na estimativa de vazões com tempos de recorrência muito elevados?

2) Qual a magnitude destes erros?

3) Qual a probabilidade de que as estimativas sejam inferiores aos valores reais?

4) Qual a influência do tamanho das amostras disponíveis na precisão dos resultados?

São geradas séries sintéticas representativas dos registros de vazões diárias de alguns postos da bacia do Rio Doce, a partir dos quais a metodologia de avaliação é aplicada para comparar os métodos de máximos anuais aos métodos de séries parciais.

Naturalmente, as respostas às quatro perguntas anteriores só podem ser dadas a partir do conhecimento da população. Os resultados obtidos são, então, específicos para as populações estudadas.

Na prática, deve-se escolher um modelo a partir de evidências amostrais, pois não se conhece a população. Em geral, os testes de adequação de ajuste são usados como suporte para essa escolha, assumindo-se a hipótese de que a distribuição teórica que melhor se ajustar à amostra será também a que mais se aproximará da verdadeira população. Aproveitando as séries sintéticas disponíveis, ainda no Capítulo IV é apresentada uma metodologia simples para a verificação desta hipótese.

Finalmente, no Capítulo V são discutidos os resultados obtidos nos capítulos precedentes, e apresentadas algumas conclusões e recomendações.

CAPÍTULO II

ANÁLISE DIRETA DE FREQUÊNCIA DE CHEIAS

Foram mostradas no item 1.2 diferentes formas de se tratar um registro de vazões diárias, com o objetivo de se obter uma amostra que permitisse a análise direta de frequência de cheias. Neste capítulo, são expostas as modelagens matemáticas das séries de máximos anuais e das séries parciais, procurando ressaltar as hipóteses subjacentes aos dois casos. Estas hipóteses serão verificadas em um caso-exemplo, a partir do levantamento de diversas estatísticas das séries de máximos anuais e das séries parciais da bacia do Rio Doce. Será tratada ainda a possibilidade do aproveitamento de informações regionais para uma melhor caracterização dos dois tipos de séries. Novamente, os registros de vazões diárias da bacia do Rio Doce servirão de base para um exemplo de aplicação.

2.1 - Séries de Máximos Anuais e Séries Parciais

Existem diferentes expedientes para se extrair informações de um registro de vazões médias diárias para a análise de frequência de cheias. O mais comum consiste em considerar na análise apenas o maior pico de cheia ocorrido em cada ano hidrológico, isto é, a máxima vazão diária em cada ano ou, simplesmente, máximo anual. Uma objeção frequentemente levantada a este procedimento é que a consideração de apenas um valor de vazão a cada ano pode acarretar uma perda de informação. De fato, a segunda maior cheia de um dado ano, omitida na série de máximos anuais, po

de ser maior que vários elementos dessa série. Uma possível forma de se evitar este inconveniente é considerar, na análise de frequência de cheias, as chamadas séries parciais.

A série contendo os maiores eventos (picos ou volumes) registrados em cada cheia, cujas magnitudes ultrapassem um certo nível-base ou nível de truncamento, é denominada série parcial. Esta série envolve, a princípio, duas variáveis aleatórias:

- 1) O número de eventos ocorridos em cada ano, e
- 2) A magnitude destes eventos.

Assim, devem ser desenvolvidos métodos probabilísticos para análise das séries parciais diferentes daqueles usados nas modelagens de séries de máximos anuais. Embora uma série parcial extraída de um registro de vazões diárias possa conter um maior número de elementos que a série de máximos anuais correspondente, isto não significa, necessariamente, que seu conteúdo de informação seja maior. De fato, um mesmo fenômeno hidro-meteorológico pode dar origem a vários picos de cheia sucessivos, que podem ser dependentes, o que reduziria o conteúdo de informação da série.

O tempo de recorrência de uma vazão de certa magnitude é definido como o tempo médio entre a ocorrência de duas vazões que ultrapassem esta magnitude. Os tempos de recorrência das séries de máximos anuais e das séries parciais têm diferentes significados. No primeiro caso, o tempo de recorrência é o número médio de anos entre a ocorrência de dois máximos anuais que superem certa magnitude. No caso de séries parciais, o tempo de re-

corrência não está associado à ocorrência de máximos anuais, ou seja, o tempo médio entre a ocorrência de picos acima de um certo valor não é o mesmo que o tempo médio entre a ocorrência de máximos anuais acima deste valor.

Chow (1964) investigou teoricamente as relações entre estes dois tempos de recorrência e sua correspondente probabilidade. Seja p_p a probabilidade de uma certa vazão q ser excedida numa série parcial contendo λN picos, onde N é o número de anos da série e λ o número médio de picos por ano. Então, p_p/λ é a probabilidade de ocorrência de um pico com magnitude superior a q em um ano qualquer. A probabilidade de uma vazão q ser a maior das λ que ocorrem (em média) em um ano é:

$$p_{\lambda} = \left(1 - \frac{p_p}{\lambda} \right)^{\lambda} \quad (2.1)$$

Seja agora p_a a probabilidade de um máximo qualquer ser superior a q numa série de máximos anuais. Admitindo a aproximação

$$\left(1 - \frac{p_p}{\lambda} \right)^{\lambda} \approx e^{-p_p} \quad (2.2)$$

$$\begin{array}{l} \text{e fazendo} \\ \text{e} \end{array} \quad \left. \begin{array}{l} T_a = \frac{1}{p_a} \\ T_p = \frac{1}{p_p} \end{array} \right\} \quad (2.3)$$

onde T_a e T_p são os tempos de recorrência da vazão q na série de máximos anuais e na série parcial, respectivamente, tem-se:

$$T_p = \frac{1}{\ln T_a - \ln(T_a - 1)} \quad (2.4)$$

Esta distinção, no entanto, só é importante nos casos de pequenos tempos de recorrência (Kite, 1977).

A utilização de séries parciais em análise de frequência de vazões tem sido recomendada apenas nos estudos que envolvam pequenos tempos de recorrência (Chow, 1964; Kite, 1977; Kottegoda, 1980). Para a estimação de vazões com tempos de recorrência elevados, NERC (1975) considera desaconselhável a utilização de séries parciais, sem uma verificação cuidadosa das hipóteses subjacentes, limitando-se a sugerir seu uso para estimar a média dos máximos anuais nos casos em que os registros disponíveis forem curtos (menos de 10 anos). Entretanto, não se pode deixar de mencionar algumas vantagens potenciais do uso de séries parciais, como a possibilidade de amostras com maior número de elementos (o que não é, necessariamente, garantia de melhores estimativas) e a possibilidade de se considerar explicitamente a ocorrência de máximos anuais provenientes de diferentes distribuições (sazonalidade).

2.2 - Modelos de Máximos Anuais

A determinação de vazões com tempo de recorrência T anos a partir da análise de séries de máximos anuais tem a característica de apresentar uma modelagem matemática simples, uma

vez que se dispõe de uma amostra tirada diretamente da população que se quer estudar: vazões máximas anuais. Entretanto, existem algumas hipóteses implícitas nos modelos de análise de frequência de máximos anuais que merecem ser olhadas com atenção:

- 1) Os dados a serem analisados descrevem um evento aleatório independente.
- 2) Os fenômenos naturais envolvidos no processo são estacionários com respeito ao tempo.
- 3) Os parâmetros da população podem ser estimados a partir da amostra.

A estacionariedade e a independência das séries de máximos anuais permitem que sua análise estatística e o tratamento probabilístico de vazões futuras (que, em última análise, é o que se deseja) possam ser feitos de forma mais confiável e simples. Em geral, o único prejuízo em se tratar como aleatória uma série dependente é que a dependência atua como que diminuindo o tamanho da amostra, pois um determinado elemento da mesma pode não acrescentar informação alguma às fornecidas pelos demais. Isto faz com que a variância dos estimadores aumente, diminuindo o nível de confiança das estimativas. Costuma-se definir o início do ano hidrológico como sendo o início do período chuvoso da região. Assim, consegue-se evitar dois máximos anuais provenientes de um mesmo período de cheias, pois, neste caso, entre dois máximos consecutivos sempre existirá um período de estiagem.

A não-estacionariedade da série de máximos anuais pode causar prejuízos maiores, pois modifica não só a variância dos

estimadores, como leva a estimativas tendenciosas da média e da assimetria da distribuição (Kite, 1977). Se a tendência histórica pudesse ser determinada, a abordagem de análise de frequência poderia ser feita subtraindo-se dos valores da amostra o termo de terminístico e analisando a nova amostra sem tendência. Entretanto, a extensão dos registros de vazões disponíveis é, em geral, pequena para que uma possível tendência histórica possa ser perfei-tamente identificada. Na melhor das hipóteses, consegue-se detectar esta tendência através de testes estatísticos adequados: neste caso, as técnicas que serão aqui tratadas não se aplicam.

Poder-se-ia argumentar que, como a natureza está em constante evolução, não existe a almejada série estacionária, e, portanto, toda análise ou tratamento probabilístico que se baseas se nesta hipótese seria, por princípio, errada. Em contrapartida, existem situações em que o grau de não-estacionariedade é suficientemente reduzido para que se possa, sem grandes problemas, simplesmente ignorá-lo. Este é o caso de uma bacia hidrográfica cuja evolução natural se processa através de mudanças graduais, em escala geológica. Os desvios desta situação mais comumente encontrados (Damázio e Kelman, 1983) são as mudanças bruscas causadas por acidentes naturais como incêndios em florestas, ou por interferên cia repentina do homem no ciclo hidrológico (Exemplo: Início de operação de reservatórios) e as mudanças graduais velozes (recuperação da floresta após o incêndio, urbanização). Além disso, uma série de vazões extremas pode apresentar não-estacionariedade devido a alguma inconsistência dos dados causada, por exemplo, por mudança de régua ou do local de medição.

Damazio e Kelman (1983) apresentam um estudo da aleatoriedade e da estacionariedade das séries de máximos anuais de várias durações aplicado aos postos das bacias dos Rios Doce, Paraíba do Sul e Paranapanema. O número de séries para as quais foi rejeitada a hipótese de estacionariedade a um nível de significância de 1% (5 séries num total de 36 na bacia do rio Doce) indicou não ser prudente a utilização desta hipótese sem uma confirmação através de testes.

Se as hipóteses de aleatoriedade e de estacionariedade não puderem ser rejeitadas, pode-se, simplesmente, ajustar uma distribuição de probabilidades adequada, cuja função de distribuição será denotada por $F(x)$, aos N máximos anuais correspondentes a N anos de registro. É preciso, ainda, que seja possível a estimação dos parâmetros desta distribuição a partir da amostra disponível. Desta forma, a vazão com tempo de recorrência T anos (ou quantil de T anos) é definida como:

$$q(T) = x, \text{ tal que } P[Q > x] = 1/T$$

ou seja: $q(T) = x, \text{ tal que } F(x) = 1 - 1/T$

De maneira geral, a escolha da distribuição $F(x)$ mais adequada para representar a população X é feita a partir de testes de adequação de ajuste. Admite-se que uma distribuição pode representar de forma adequada uma certa população se a hipótese de que a amostra disponível foi sorteada desta distribuição não puder ser rejeitada por estatísticos apropriados. Uma boa descrição dos testes de adequação de ajuste mais utilizados pelos

hidrólogos, acompanhada de uma análise crítica dos mesmos, pode ser encontrada em Gomide e Kaviski (1981). Pinheiro (1982) apresenta aplicações de testes de adequação de ajuste a inúmeras séries de máximos anuais de postos das regiões sul e sudeste do Brasil.

A medida que se quer incorporar aos modelos de máximos anuais novas evidências empíricas, é preciso procurar distribuições de probabilidades capazes de reproduzir estas peculiaridades. Assim, as distribuições de uso mais comum em análise de frequência vão se tornando mais e mais completas, à medida que novas características das séries de extremos são verificadas. Por exemplo, a verificação de que as distribuições empíricas dos dados de vazões extremas se mostravam assimétricas levou os hidrólogos do início do século a abandonarem os resultados estabelecidos a partir da teoria dos erros e a procurarem, então, distribuições de probabilidades capazes de reproduzirem estas assimetrias amostrais (Chow, 1964). A partir deste primeiro avanço, diversas outras constatações empíricas colocaram em evidência distribuições tais como log-normal, Pearson tipo 3, Log-Pearson tipo 3 e, mais recentemente, as distribuições Wakeby e Lambda.

Entretanto, nem todas as distribuições usualmente empregadas na análise de frequência de vazões extremas partem de verificações empíricas. Um exemplo deste fato é a explicação teórica dada por Chow (1963) para o uso da distribuição log-normal para representar os máximos anuais, partindo da hipótese e de que o valor de uma variável hidrológica é determinado pelo produto de diversos fatores devidos a fenômenos mais elementares. Desta forma, o logarítmo deste valor é igual à soma dos logarítmos dos vá-

rios fatores, apresentando uma distribuição normal. Fisher e Tippet (1928) mostraram que existem apenas três casos estáveis de distribuições de probabilidades para o máximo de N variáveis aleatórias independentes e identicamente distribuídas, se N for um número muito grande. Gumbel (1958, pág. 238), discutindo a aplicação destas três distribuições às vazões máximas anuais, argumentou que, das três hipóteses assumidas, ou seja,

- 1) A distribuição das vazões diárias é do tipo exponencial.
- 2) $N = 365$ é um número suficientemente grande de observações.
- 3) As observações diárias são independentes,

as duas primeiras não podem ser verificadas, uma vez que a distribuição das vazões diárias não é conhecida, e a terceira é certamente falsa, o que acarreta um número de observações independentes a cada ano inferior a 365. Além deste inconveniente, não há como deduzir teoricamente qual dos três tipos de distribuição de extremos deve ser utilizado (NERC, 1975), sendo necessária considerável evidência empírica para demonstrar a aplicabilidade deste ou daquele tipo. Assim, não obstante os trabalhos teóricos já desenvolvidos, a escolha da distribuição a ser usada acaba sendo feita empiricamente.

Uma análise cuidadosa dos métodos de análise de frequência de cheias baseados nas séries de máximos anuais pode ser encontrada em Pinheiro (1982). Neste trabalho são apresentadas di

versas aplicações destes métodos na determinação de quantis com diferentes tempos de recorrência para bacia hidrográficas situadas nas regiões sul e sudeste do Brasil.

2.3 - Métodos de Séries Parciais

2.3.1 - Modelagem Matemática

Em uma série parcial, cada elemento corresponde à magnitude de uma cheia ocorrida num certo ano hidrológico, ou seja, à maior vazão ocorrida durante esta cheia, sendo a vazão máxima anual igual à magnitude da maior cheia ocorrida neste ano. Como o número de cheias que ocorre em cada ano é uma variável aleatória, o estudo da distribuição de probabilidades dos máximos anuais a partir das séries parciais é equivalente ao estudo da distribuição do máximo valor de uma série cujo número de termos é também uma variável aleatória.

Borgman (1963) e Shane e Lynn (1964) admitiram a hipótese de que o número de picos ou cheias que ocorre a cada ano obedece às leis de um processo Poisson independente do tempo. As bases de tal processo podem ser resumidas em:

- 1) A probabilidade da ocorrência de um evento num certo intervalo de tempo é proporcional ao tamanho do intervalo;
- 2) Os números de ocorrências em intervalos que não se superpõem são independentes;

- 3) Dois eventos não podem ocorrer simultaneamente;
- 4) O número de ocorrências em um dado intervalo é independente da situação deste intervalo no tempo.

Kirby (1969) propõe um modelo para séries parciais, considerando uma hidrôgrafa como uma sequência de cheias, aproximadamente instantâneas, separadas por períodos relativamente longos de vazões baixas. Segundo este modelo, a ocorrência de picos acima de um nível de truncamento poderia ser considerada como um sucesso numa sequência de sorteios do tipo Bernoulli aleatoriamente espaçados no tempo. A partir desta formulação, foi possível mostrar que, para uma probabilidade de sucesso suficientemente pequena, a distribuição de probabilidades do tempo entre eventos e do número de eventos em um dado intervalo de tempo segue as leis de um processo Poisson. Isto justificaria a hipótese assumida anteriormente por Borgman e por Shane e Lynn de que a variável aleatória "número de picos em um ano" tem uma distribuição próxima da de Poisson, pelo menos se forem escolhidos níveis de truncamento suficientemente elevados.

Os modelos descritos até aqui têm o inconveniente de não levarem em conta a variável aleatória "magnitude dos picos" em sua formulação. Entretanto, Shane e Lynn (1964) já ressaltavam a excelente qualidade do ajuste da distribuição exponencial às séries parciais. A explicação deste fato é que as séries parciais são, na verdade, uma amostra de uma população cuja função de distribuição corresponde à cauda da função de distribuição de uma po

pulação maior. Neste caso, se a distribuição de probabilidade das vazões diárias tiver cauda exponencial, a distribuição dos picos deverá ser exponencial.

Todorovic (1970) derivou uma expressão analítica para a distribuição de probabilidades da variável aleatória "máximo evento de uma amostra contendo um número aleatório de elementos". Ora, a máxima vazão anual é igual ao maior pico de cheia registrado no ano, sendo o número de cheias, também, uma variável aleatória. Esta expressão de Todorovic deu origem a vários modelos para análise de frequência de cheias, a partir das séries parciais. Para melhor compreensão destes modelos, seguem-se uma rápida exposição das hipóteses feitas e a derivação da expressão geral.

Seja $N(t)$ o número de picos que ocorre num certo intervalo $[0, t]$ e seja X_i , $i = 1, 2, \dots, N(t)$ a magnitude de cada um deles. Tomam-se como hipóteses:

H1. X_i , $i = 1, 2, \dots, N(t)$ é uma sequência de variáveis aleatórias independentes e com distribuição de probabilidades conhecida;

$$P \{X_i \leq x\} = H_i(x) \quad (2.5)$$

H2. Os valores da sequência X_i , $i = 1, 2, \dots, N(t)$ são independentes do número de termos da mesma, $N(t)$.

A distribuição de probabilidades do maior evento ocorrido no intervalo $[0, t]$ é dada por:

$$F_t(x) = P\{[\max X_i, i = 1, 2, \dots, N(t)] \leq x\} =$$

$$= P\left\{\bigcap_{i=1}^{N(t)} X_i \leq x\right\} \quad (2.6)$$

Com o auxílio das hipóteses H1 e H2 acima, pode-se concluir:

$$F_t(x) = P\{N(t)=0\} + \sum_{n=1}^{\infty} \left[\prod_{k=1}^n H_k(x) P\{N(t) = n\} \right] \quad (2.7)$$

Se o intervalo $[0, t]$ for igual a um ano, esta é uma expressão exata para a distribuição dos máximos anuais desde que, naturalmente, as hipóteses H1 e H2 sejam aceitáveis. Além disso, esta expressão permite uma generalização para que se leve em conta a sazonalidade no processo de ocorrência dos picos ao longo do ano. Assim, um estudo adequado das variáveis aleatórias $N(t)$ e X_i pode tornar o uso de séries parciais potencialmente mais atraente para análise de frequência de cheias que o uso de séries de máximos anuais.

Admitindo-se a hipótese de que o número de picos a cada ano tem distribuição Poisson (Borgman, 1963; Shane e Lynn, 1964; Kirby, 1969), e que as variáveis $X_i, i = 1, 2, \dots$ são identicamente distribuídas e independentes tem-se:

$$P\{N(\text{um ano}) = n\} = \frac{e^{-\lambda} \lambda^n}{n!} \quad (2.8)$$

$$H_i(x) = H(x), \quad \forall i = 1, 2, \dots \quad (2.9)$$

Substituindo as equações 2.8 e 2.9 na equação 2.7 pode-se obter a seguinte expressão para a distribuição dos máximos anuais (Todorovic e Shen, 1976):

$$F(x) = \exp\{-\lambda[1 - H(x)]\} \quad (2.10)$$

onde λ é o valor esperado do número de picos em um ano. Esta expressão foi inicialmente estabelecida por Epstein (1949) e utilizada posteriormente por Todorovic e Zelenhasic (1970).

Seja $Q(T) = q_T$ o máximo anual com tempo de recorrência T anos, ou seja:

$$F(q_T) = 1 - \frac{1}{T} \quad (2.11)$$

Substituindo esta equação na equação 2.10 e tirando os logaritmos, tem-se:

$$\ln\left[1 - \frac{1}{T}\right] = -\lambda[1 - H(q_T)] \quad (2.12)$$

Sendo a probabilidade de um certo pico de cheia ser inferior a q_T dada por:

$$H(q_T) = 1 + \frac{1}{\lambda} \ln\left[1 - \frac{1}{T}\right] \quad (2.13)$$

Se em um ano ocorrem em média λ picos, o pico com tempo de recorrência T_p anos deve ser ultrapassado, em média, uma vez a cada λT_p picos, ou seja:

$$H(q_T) = 1 - \frac{1}{\lambda T_p} \quad (2.14)$$

A substituição desta equação na equação 2.13 fornece

$$1 - \frac{1}{\lambda T_p} = 1 + \frac{1}{\lambda} \ln \left(1 - \frac{1}{T} \right) \quad (2.15)$$

o que permite estabelecer a seguinte relação entre os tempos de recorrência dos máximos anuais e das séries parciais:

$$T_p = - \frac{1}{\ln \left(1 - \frac{1}{T} \right)} \quad (2.16)$$

Este resultado é idêntico ao obtido por Kite (1977) e já apresentado no item 2.2 (equação 2.4).

A expressão 2.10 pode ser simplificada, se for aceita a hipótese de que as variáveis X_i são exponencialmente distribuídas, conforme sugestão de Shane e Lynn (1964). Neste caso,

$$H(x) = 1 - \exp[-\beta(x - x_0)], \quad (2.17)$$

resultando

$$F(x) = \exp\{-\lambda \exp[-\beta(x - x_0)]\} \quad (2.18)$$

Esta expressão é uma dupla exponencial, do mesmo modo que a expressão da distribuição de Gumbel, só que, ao contrário desta última, não foi obtida a partir de considerações assintóticas (Gupta et Alii, 1976). É importante notar que a sugestão de Shane e Lynn

se baseia em duas hipóteses:

- 1) O nível de truncamento é muito elevado;
- 2) A distribuição das vazões diárias apresenta comportamento exponencial na cauda direita.

A equação 2.18 pode ser tomada como a própria expressão da distribuição de Gumbel, apenas apresentando um outro significado para os parâmetros, ou seja, uma outra maneira de estimá-los. Analogamente, pode-se verificar que, se as magnitudes dos picos X_i têm distribuição de probabilidades Pareto, ou seja:

$$H(x) = 1 - \left(\frac{k}{x} \right)^\alpha \quad (2.19)$$

Então a distribuição dos máximos anuais associada será dada por:

$$F(x) = \exp \left[-\lambda \left(\frac{k}{x} \right)^\alpha \right] \quad (2.20)$$

que é a expressão da distribuição de extremos tipo II.

Zelenhasic (1970) estudou o modelo de Todorovic (1970), considerando duas hipóteses adicionais:

- H3. A variável $N(t)$ segue um processo Poisson dependente do tempo, ou seja:

$$P\{N(t) = n\} = e^{-A(t)} \frac{[A(t)]^n}{n!} \quad (2.21)$$

sendo $A(t)$ o valor esperado do número de picos no intervalo $[0, t]$.

H4: As variáveis X_i , $i = 1, 2, \dots$ são identicamente distribuídas com distribuição de probabilidades $H(x)$.

O modelo proposto permite considerar os efeitos da sazonalidade na taxa de ocorrência de picos ao longo do ano, o que é particularmente interessante quando se quer determinar vazões com tempos de recorrência pequenos (NERC, 1975).

Para considerar a sazonalidade de forma mais geral, Todorovic e Rousselle (1971) apresentaram um modelo semelhante ao de Zelenhasic, considerando a distribuição das magnitudes dos picos como exponencial, com os parâmetros variando ao longo do ano. A aplicação destes dois modelos é feita dividindo-se o ano em estações de pequena duração e estudando as variáveis $N(t)$ e X_i em cada uma delas. Todorovic e Woolhiser (1972), partindo das mesmas idéias, estudaram a sequência de variáveis aleatórias T_i , $i = 1, 2, \dots$ sendo T_i a data de ocorrência do i -ésimo pico.

Finalmente, em 1976, Gupta, Duckstein e Peebles obtiveram uma expressão analítica para a distribuição conjunta do maior pico de cheia e sua data de ocorrência, a partir das hipóteses H1 e H2 já expostas. Este resultado, que é uma generalização das expressões obtidas por Todorovic e Zelenhasic (1970) e por Todorovic e Woolhiser (1972), possibilita uma descrição mais completa do processo das séries parciais.

Até aqui, a variável aleatória "magnitude das cheias" tem sido associada aos picos das mesmas. Entretanto, a magnitude

da cheia pode ser associada à sua duração, seu volume ou qualquer outra grandeza que se queira estudar. Todorovic (1978) estudou a distribuição de probabilidades dos volumes de cheias a partir das séries parciais, considerando o volume de uma cheia, V_i , acima de um nível de truncamento, x_0 , dado pela expressão:

$$V_i = \frac{(X_i - x_0) T_i}{2} \quad (2.22)$$

sendo X_i e T_i o pico e a duração da i -ésima cheia, respectivamente. Este estudo já envolve três variáveis aleatórias: o número de cheias por ano, a duração e o pico de cada uma.

Modelos mais sofisticados para estudo de frequência de cheias a partir de séries parciais continuam sendo propostos (North, 1980; Kavvas, 1982). Estes modelos permitem uma caracterização detalhada do processo de ocorrência de cheias, exigindo, em contrapartida, uma quantidade cada vez maior de informações. O modelo de Kavvas (1982), por exemplo, considera a ocorrência de picos de cheias como produto de dois processos estocásticos encadeados. O primeiro deles descreve a ocorrência das precipitações na bacia hidrográfica e segue as leis de um processo Poisson. Cada precipitação - ou "evento primário" - dá origem a uma série de picos de cheias - ou "eventos secundários" que é estudada em detalhe (é possível modelar inclusive a estrutura de dependência existente nesta série de eventos secundários). Os dois processos (precipitações e picos associados) podem considerar a sazonalidade, tornando este modelo muito atraente do ponto de vista conceitual.

Quando se quer determinar vazões com tempos de recorrência muito elevados, é possível, entretanto, tratar o proces

so de séries parciais de forma mais superficial (NERC, 1975). Com efeito, Bardsley e Manly (1979) apresentaram um método para estimar o erro cometido ao se substituir a variável aleatória $N(t)$ pelo seu valor esperado, isto é, ao se considerar que em todos os anos ocorre um número constante de cheias: verificaram que este erro decresce, tendendo assintoticamente a zero, à medida que o valor esperado do número de eventos aumenta. Este resultado permite que se ignore a distribuição de probabilidade de $N(t)$ quando t for grande, considerando apenas que seu valor esperado é dado por:

$$E\{N(T_{\text{anos}})\} = \lambda T \quad (2.23)$$

sendo λ o valor esperado do número de eventos em um ano.

NERC (1975) sugeriu um método simples para a determinação da vazão com um tempo de recorrência qualquer T , a partir das séries parciais, que pode ser resumido nos seguintes passos:

- 1) Toma-se uma série parcial contendo λ picos por ano, em média.
- 2) Ajusta-se uma distribuição de probabilidades conveniente $H(x)$ à magnitude dos picos.
- 3) O valor esperado do número de picos em T anos é λT .
- 4) A vazão com tempo de recorrência T anos é a que tem uma probabilidade de ser ultrapassada em um ano qualquer igual a $1/\lambda T$, ou seja:

$$1 - H(q_T) = \frac{1}{\lambda T} \quad (2.24)$$

Nesta equação q_T é a vazão com tempo de recorrência T anos a ser determinada.

Este método é de aplicação tão simples quanto os métodos de máximos anuais, restando saber se conduz a melhores resultados ou não. Como o número de elementos de uma série parcial é maior que o da série de máximos anuais correspondente, é de se esperar que a primeira tenha um maior conteúdo de informação que esta última, desde que os picos não sejam dependentes entre si.

Um estudo comparativo entre as estimativas de vazões com diferentes tempos de recorrência a partir de séries de máximos anuais e de séries parciais é apresentado em NERC (1975). Este estudo é teórico e parte das mesmas hipóteses que conduziram à expressão 2.10, já discutidas anteriormente. O critério de comparação utilizado foi o da variância das estimativas, e a conclusão foi que para amostras pequenas (até 20 anos de registro) o método de séries parciais apresenta resultados menos dispersos, desde que λ seja superior a 1,67. Além disso, para quantis de pequeno tempo de recorrência ($T \sim 5$ anos) este método é incondicionalmente melhor. Yevjevich e Taesombut (1979), chegaram a conclusões semelhantes, partindo das mesmas hipóteses, e confirmaram estas conclusões com o auxílio da geração de séries sintéticas de vazões diárias. Já para amostras maiores (registros com mais de 50 anos), estes autores observam que as previsões de quantis a partir das séries de máximos anuais fornecem resultados menos dispersos. NERC (1975) explicou esta aparente contradição sugerindo

que um máximo anual fornece na realidade duas informações - a sua magnitude e o fato de que este foi o maior valor registrado no ano -, além de ser a variável diretamente implicada no estudo.

É importante voltar a frisar, neste ponto, que os modelos mencionados aqui se baseiam na hipótese de independência entre os picos, entre outras. Assim como nas séries de máximos anuais costuma-se trabalhar com o ano hidrológico, ao invés do ano civil, para evitar a ocorrência de dois máximos anuais consecutivos provenientes da mesma enchente, existem critérios para a obtenção de séries parciais com menor grau de não aleatoriedade. Dois destes critérios são discutidos a seguir.

2.3.2 - Crítérios Para a Obtenção das Séries Parciais

Para a perfeita identificação das séries parciais num dado registro de vazões diárias é preciso isolar todas as cheias, o que só é viável se são disponíveis informações acerca dos fenômenos hidro-meteorológicos, dos quais elas resultam. Na falta destas informações, têm sido propostos critérios para a obtenção das séries parciais, utilizando-se apenas os dados de vazões diárias. Dois destes critérios serão discutidos aqui.

O primeiro critério (critério a) é baseado em se definir uma cheia como o conjunto de vazões que ultrapassam uma certa magnitude q_0 (denominada nível-base ou nível de truncamento da série) durante um intervalo de tempo contínuo. Este intervalo de tempo (duração da cheia) é chamado período de excesso. De cada cheia é considerada, por exemplo, a máxima vazão ocorrida, (ou o máximo excesso) como sendo a magnitude da cheia. A série con-

tendo as magnitudes das cheias, ou os picos, é denominada série parcial. A Figura 2.1 ilustra este procedimento, sendo a série parcial composta dos elementos X_1 a X_5 . Pode-se observar que, dependendo do nível de truncamento adotado, duas cheias podem ser consideradas como apenas uma, de maior duração. Este fato é esclarecido pelas Figuras 2.1 e 2.2: para o nível de truncamento q_0 (Figura 2.1), tem-se uma série parcial com 5 picos, X_i , e para o nível-base q_1 , maior que q_0 (Figura 2.2), a nova série tem 6 picos, Y_i , tendo sido a última cheia para o nível q_0 desmembrada em duas para o nível q_1 . Apesar deste inconveniente, esta metodologia auxilia a obtenção de séries com alto grau de aleatoriedade, a partir de uma escolha adequada do nível de truncamento.

O segundo critério (critério b) consiste em se considerar qualquer máximo local da hidrôgrafa como sendo um pico de cheia, definindo-se a série parcial como o conjunto de todos os picos assim obtidos cujos valores ultrapassem um certo nível de truncamento q_0 . Desta forma as cheias são definidas independentemente do nível de truncamento, sendo este usado apenas para selecionar as vazões cujos valores se situam dentro da faixa que apresenta interesse na análise de extremos. O inconveniente deste procedimento é que a série parcial assim obtida pode conter sucessivos picos provenientes de uma mesma cheia e, conseqüentemente, portanto algum grau de dependência. Pode-se observar na Figura 2.3 uma série parcial obtida de acordo com este segundo critério, onde os picos Z_7 , Z_8 e Z_9 são, aparentemente, provenientes de uma mesma cheia. NERC (1975; pág. 16) propõe uma medida para evitar este tipo de problema e tentar garantir a aleatoriedade da série parcial: dois máximos locais são considerados dois picos de cheias apenas quando:

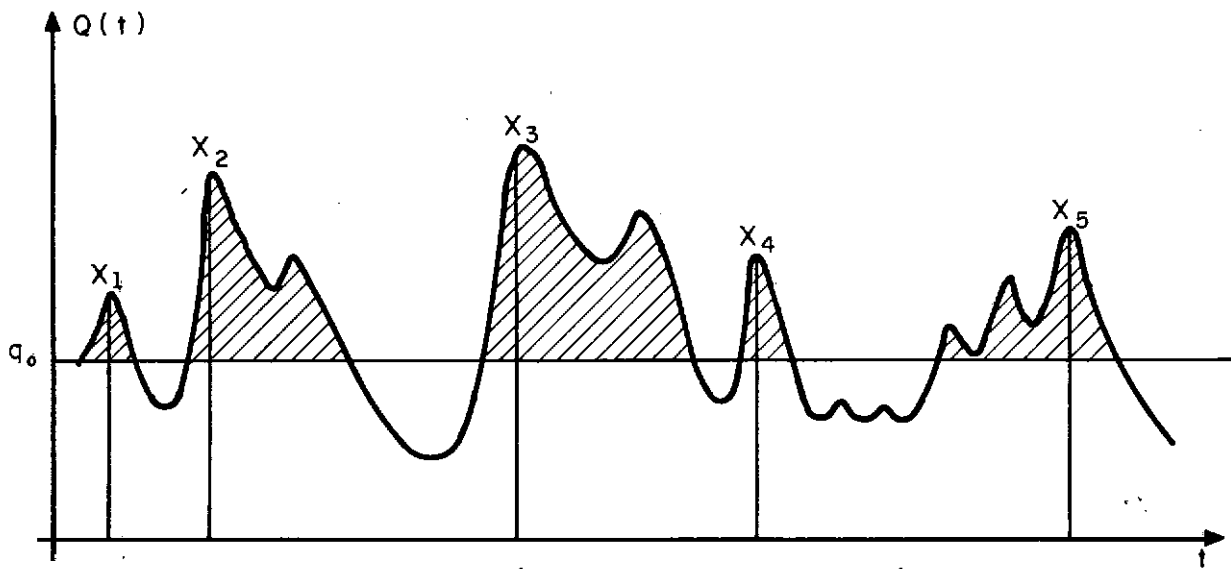


FIGURA 2.1 — OBTENÇÃO DAS SÉRIES PARCIAIS PELO CRITÉRIO A.

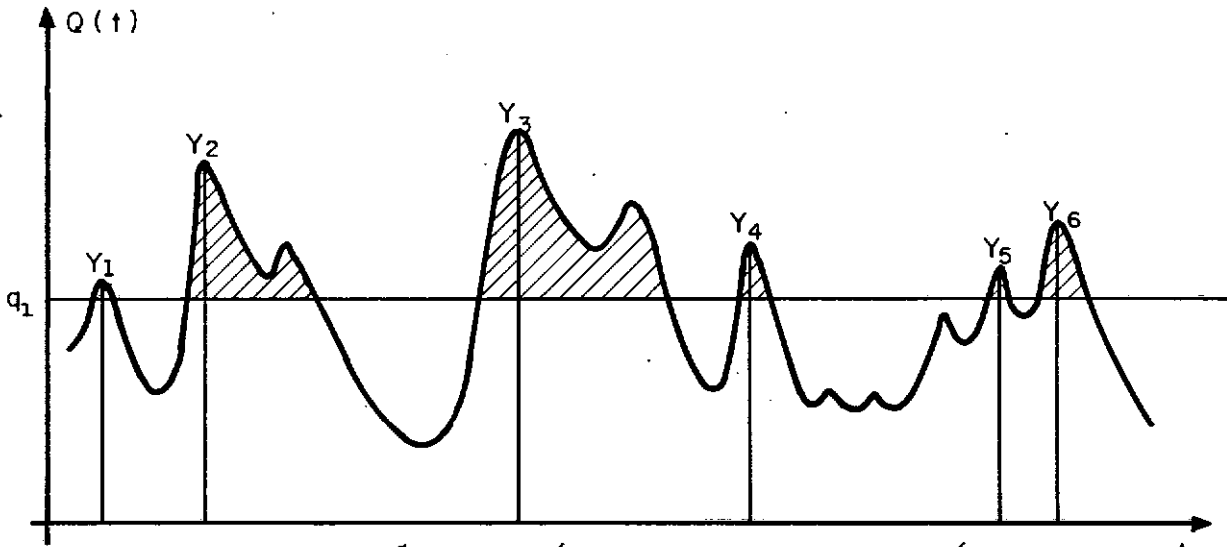


FIGURA 2.2 — OBTENÇÃO DAS SÉRIES PARCIAIS PELO CRITÉRIO A.
NÍVEL DE TRUNCAMENTO $q_1 > q_2$

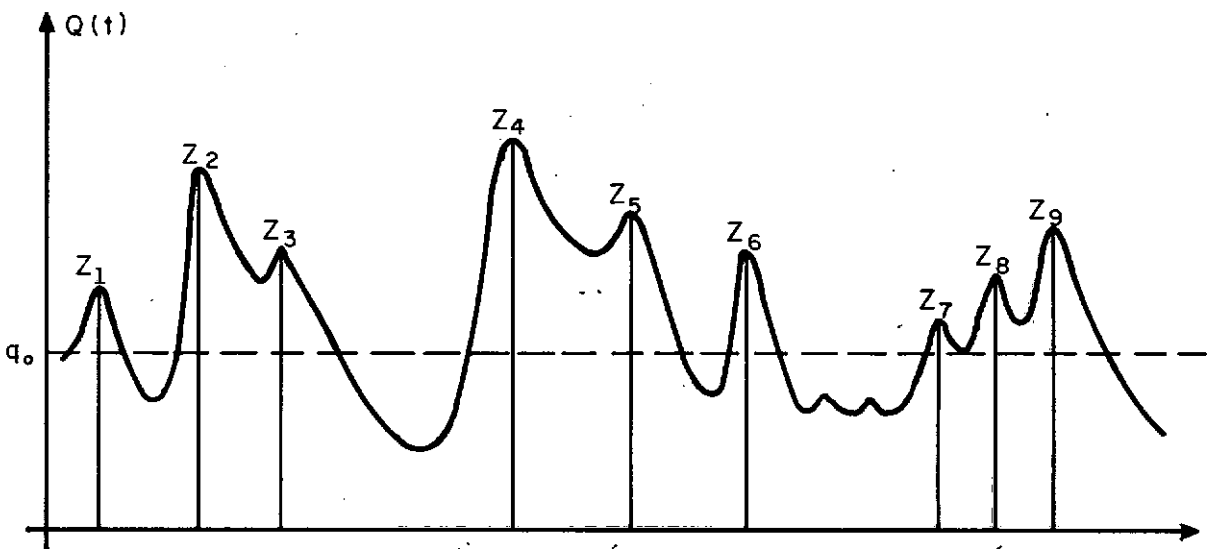


FIGURA 2.3 — OBTENÇÃO DAS SÉRIES PARCIAIS PELO CRITÉRIO B.

- 1) Estão separados por tempo superior a três vezes o tempo de concentração da bacia.
- 2) A menor vazão ocorrida entre os dois é inferior a dois-terços da magnitude do primeiro.

Caso estas duas condições não sejam satisfeitas, considera-se como pico de cheia apenas o maior dos dois. Este procedimento tem o inconveniente de exigir informações adicionais para a determinação do tempo de concentração da bacia. Water Resources Council (1976) propõe um critério semelhante, exigindo que dois picos sucessivos estejam separados por

$$\Delta t = 5 \ln A \text{ dias} \quad (2.25)$$

onde A é a área da bacia expressa em milhas quadradas. Além disso, a vazão entre os dois deve cair a 75% da magnitude do menor deles.

Ashkar e Rousselle (1983) apresentam algumas reservas a este tipo de procedimento, considerando que estas restrições impostas ao intervalo de tempo entre os picos interferem com as hipóteses básicas do processo Poisson, comumente adotadas na modelagem da variável aleatória "número de picos num dado período". De fato, a hipótese de que o número de picos por ano segue um processo Poisson tem como consequência que o intervalo entre os picos tem uma distribuição exponencial. As restrições impostas a estes intervalos truncam esta distribuição, alterando o processo de ocorrência de picos ao longo do ano. Por isso, estes autores sugerem que antes de se adotar tais restrições devem ser investi-

gadas outras medidas para a redução da correlação, como, por exemplo, a elevação do nível de truncamento.

Na escolha do critério mais adequado para a obtenção das séries parciais, deve-se ter em mente que, se as séries obtidas forem aleatórias, quanto maior o número de picos em cada uma maior será seu conteúdo de informação. Assim, um bom critério para obtenção das séries parciais deve fornecer séries aleatórias e com grande taxa anual de picos (número médio de picos por ano).

2.3.3 - A Aleatoriedade das Séries Parciais

Apresenta-se a seguir uma metodologia para a avaliação do grau de aleatoriedade das séries parciais. Esta metodologia é aplicada às séries parciais da bacia do Rio Doce, obtidas segundo os dois critérios de obtenção descritos no item anterior, com o objetivo de se ter uma idéia do desempenho dos mesmos.

A aleatoriedade das séries parciais, além de ter grande influência no seu conteúdo de informação, é uma hipótese assumida pela maioria dos modelos citados no item 2.3.1. Entretanto, o armazenamento de água pela bacia hidrográfica, aliado à grande proximidade das chuvas nos períodos úmidos do ano, faz com que uma cheia qualquer exerça grande influência sobre a seguinte, criando uma dependência entre as duas. Isto é ainda agravado pelo fato de que um mesmo fenômeno meteorológico pode dar origem a várias cheias, logicamente dependentes. Assim, pode-se partir da hipótese de que não existem séries parciais totalmente aleatórias (Damazio e Kelman, 1983), sendo possível, sim, a existência de sê

ries com um maior ou menor grau de aleatoriedade, que pode ser quantificado através de testes clássicos de hipóteses.

O Teste de Kendall, para correlação entre duas variáveis aleatórias, apresenta a característica interessante de que sua estatística tem uma distribuição de probabilidades que independe das populações que forneceram as amostras: é, portanto, um teste não-paramétrico. Este teste é descrito sumariamente a seguir.

Seja uma amostra contendo N duplas de observações associadas, de duas populações infinitas X e Y, cujos pares (x_i, y_i) estão ordenados por ordem crescente dos valores da variável X, ou seja:

$$\left[(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) \right]$$

com

$$x_1 < x_2 < \dots < x_n$$

Ao ser feita esta ordenação dos pares por ordem crescente dos valores de X, a série (Y_1, Y_2, \dots, Y_n) pode se apresentar como qualquer das $n!$ permutações possíveis, desde que os sorteios da população Y tenham sido feitos de forma independente dos sorteios concomitantes da população X. Se, no entanto, os sorteios concomitantes das populações X e Y são monotonamente dependentes (por exemplo, linearmente dependentes), as permutações dos Y's que formarem séries monótonas serão mais prováveis. O teste é feito da seguinte maneira:

1) Calcula-se:

$$I = \sum_{i=1}^{n-1} S_i \quad (2.26)$$

sendo S_i o número de Y_j maiores que Y_i , com $i < j \leq n$.

2) Calcula-se:

$$T = \sum_{i=1}^{n-1} t_i \quad (2.27)$$

sendo t_i o número de Y_j menores que Y_i , com $i < j \leq n$.

3) A estatística do teste é dada por:

$$S = T - I \quad (2.28)$$

A hipótese nula é a da independência entre as séries x_i e y_i . Se a hipótese nula é verdadeira, S deve ser próximo de zero. Se existe alguma tendência monótona, S tende para seu valor máximo

$$S_{\max} = \frac{n}{2} (n - 1) \quad (2.29)$$

ou mínimo

$$S_{\min} = - \frac{n}{2} (n - 1) \quad (2.30)$$

Bradley (1968) fornece os valores críticos de S para amostras contendo de 4 até 40 elementos e níveis de significância, α , iguais a 0,5%, 1%, 2,5%, 5% e 10%. Para amostras com

mais de 40 elementos deve-se usar a estatística:

$$S = \frac{S - 1}{\left[\frac{n(n-1)(2n+5)}{18} \right]^{0,5}} \quad (2.31)$$

sendo, neste caso, os valores críticos obtidos da distribuição normal padrão.

O teste de Kendall é muito sensível à persistência, podendo também detectar sinais de periodicidade.

Este teste foi aplicado às séries parciais da bacia do Rio Doce, com o objetivo de quantificar o seu grau de aleatoriedade e também verificar o desempenho dos dois critérios de obtenção das mesmas no que tange à sua capacidade de fornecer séries aleatórias. Foram utilizados os dados dos postos fluviométricos listados na Tabela 2.1. Estes postos foram selecionados por Damazio e Kelman (1983), devido ao pequeno grau de dependência e à estacionariedade verificados nas suas séries de máximos anuais. Naturalmente, se as séries de máximos anuais destes postos podem ser consideradas estacionárias, não há motivos para se esperar o contrário de suas séries parciais, o mesmo não acontecendo com respeito à aleatoriedade. Neste caso, o efeito de persistência é localizado dentro do ano hidrológico, e a aleatoriedade das séries de máximos anuais não garante a aleatoriedade das séries parciais correspondentes. Os testes foram aplicados da seguinte forma:

ORDEM	CÓDIGO DNAEE	N O M E D O P O S T O	ÁREA DE DRENAGEM	PERÍODO DE OBSERVAÇÃO	Nº DE ANOS
01	56010000	Rio Piranga em Ponte do São Lourenço	540 km ²	1938-1965	21
02	56028000	Rio Piranga em Piranga	1415 km ²	1938-1973	31
03	56050000	Rio Chopotô em Alto Rio Doce	333 km ²	1938-1965	17
04	56075000	Rio Piranga em Porto Firme	4171 km ²	1938-1973	34
05	56090000	Rio Turvo Limpo em Fazenda Varginha	331 km ²	1941-1973	28
06	56110000	Rio Piranga em Ponte Nova	6197 km ²	1939-1973	34
07	56145000	Ribeirão do Carmo em São Caetano de Mariana	213 km ²	1930-1965	32
08	56148000	Ribeirão do Carmo em Furquim	303 km ²	1934-1972	26
09	56240000	Rio Gualaxo do Sul em Fazenda Paraíso	843 km ²	1941-1973	30
10	56335000	Rio do Carmo em Acaiaca	1301 km ²	1940-1973	29
11	56337000	Rio Gualaxo do Norte em Fazenda Ocidente	523 km ²	1938-1973	20
12	56485000	Rio Matipô em Raul Soares	1306 km ²	1936-1973	34
13	56500000	Rio Santana em Abre Campo	268 km ²	1939-1973	30
14	56565000	Ribeirão do Sacramento em Bom Jesus do Galho	307 km ²	1941-1973	29
15	56670000	Rio do Peixe em Fazenda Olaria	527 km ²	1941-1965	15
16	56720000	Rio Doce em Cachoeira Escura	24245 km ²	1939-1973	25

(Cont.)

(Cont.)

ORDEM	CÓDIGO DNAEE	N O M E D O P O S T O	ÁREA DE DRENAGEM	PERÍODO DE OBSERVAÇÃO	Nº DE ANOS
17	56925000	Rio Caratinga em Caratinga	160 km ²	1938-1965	27
18	56940000	Rio Cuieté em Barra do Cuieté	3181 km ²	1938-1973	19
19	56948000	Rio Doce em Resplendor	61610 km ²	1938-1972	30
20	56967000	Rio Manhuaçu em Santana Manhuaçu	1521 km ²	1941-1965	23
21	56974000	Rio Manhuaçu em Cachoeira da Neblina	2092 km ²	1938-1965	27
22	56986000	Ribeirão São Domingos em Fazenda Boa Esperança	313 km ²	1939-1965	22
23	56988500	Rio José Pedro em Ipanema	1345 km ²	1938-1972	34
24	56989500	Rio José Pedro em Assaraí	3375 km ²	1938-1972	34
25	56990000	Rio Manhuaçu em São Sebastião da Encruzilhada	8454 km ²	1938-1972	31
26	56992000	Rio Guandú em Baixo Guandú	2162 km ²	1938-1972	32
27	56993600	Rio Santa Joana em Santa Joana	1100 km ²	1939-1968	20
28	56994000	Rio Santa Maria em Santa Maria	774 km ²	1941-1965	18
29	56994500	Rio Doce em Colatina	76615 km ²	1938-1973	28

Tabela 2.1 - Características dos Postos Fluvimétricos da Bacia do Rio Doce.

- 1) Para cada posto foram definidas séries parciais, segundo os dois critérios definidos no item 2.3.2, contendo 1, 2, 3 e 4 picos por ano, em mê dia.
- 2) Para cada série assim obtida (oito para cada posto), calculou-se a estatística de Kendall, considerando, para cada ano contendo n picos: (n maior que 1), $n-1$ duplas (x_i, x_{i+1}) com $i = 1, 2, \dots, N-1$.
- 3) Foram assinalados os casos em que a hipótese nula foi rejeitada para três níveis de significância: 1%, 5% e 10%.

Mesmo que um dos dois critérios de obtenção das séries parciais seja suficiente para fornecer séries com pequeno grau de não aleatoriedade na bacia do Rio Doce, é de se esperar que, por mero acaso, a hipótese nula seja rejeitada no teste de Kendall em alguns postos fluviométricos desta bacia. Foi feita então a seguinte consideração regional: admitindo-se a hipótese nula de aleatoriedade das séries parciais obtidas através de um certo critério e assumindo-se a hipótese adicional da independência entre as séries parciais dos diferentes postos da bacia para uma dada taxa anual de picos, o número de séries com esta taxa de picos rejeitadas no teste, a um certo nível de significância α , é uma variável aleatória com distribuição de probabilidades Binominal de média $k\alpha$ (onde k é o número de postos fluviométricos da região). A probabilidade de muitas séries serem rejeitadas cresce à medida que o critério de obtenção das mesmas não forneça séries suficientemente aleatórias.

As Tabelas 2.2 e 2.3 listam todos os postos da bacia do Rio Doce que tiveram algumas de suas séries parciais rejeitadas no teste de aleatoriedade para níveis de significância de 10%, 5% e 1%. A Tabela 2.2 se refere às séries obtidas pelo critério a (pico = maior vazão num período no qual as vazões ultrapassaram continuamente o nível de truncamento) e a Tabela 2.3 considera as séries obtidas segundo o critério b (pico = um máximo local da hidrôgrafia).

POSTO	TAXA MÉDIA ANUAL DE PICOS			
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$
56050000			5	
56075000		10		
56110000				5
56335000		10		
56500000	5			
56565000	10		5	1
56925000			10	
56948000	10			
56990000	5			
56993600		10		

Tabela 2.2 - Postos da bacia do Rio Doce cujas séries parciais foram rejeitadas pelo teste de Kendall. Séries obtidas pelo critério a.

POSTO	TAXA MÉDIA ANUAL DE PICOS			
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$
56075000	5			
56090000	1	10		
56485000			10	
56720000				5
56948000	5			5
56967000				5
56974000		5	5	1
56990000				1
56994000			10	5
56994500			10	1

Tabela 2.3 - Postos da bacia do Rio Doce cujas séries parciais foram rejeitadas pelo teste de Kendall. Séries obtidas pelo critério b.

As Tabelas 2.4 e 2.5 mostram o número de séries rejeitadas para cada taxa anual de picos, λ , para séries obtidas pelos critérios a e b, respectivamente. Aparece ainda listada nestas tabelas (entre parênteses) a probabilidade de que o número de séries rejeitadas seja maior ou igual ao número observado, sob as hipóteses de que as séries são aleatórias e independentes entre si.

α	TAXA MÉDIA ANUAL DE PICOS			
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$
1%	0 (1,000)	0 (1,000)	0 (1,000)	1 (0,182)
5%	2 (0,429)	0 (1,000)	2 (0,412)	2 (0,264)
10%	4 (0,329)	3 (0,565)	3 (0,540)	2 (0,608)
Nº DE POSTOS	29	29	28	20

Tabela 2.4 - Número de séries rejeitadas pelo teste de Kendall na bacia do Rio Doce. Séries obtidas pelo critério a.

α	TAXA MÉDIA ANUAL DE PICOS			
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$
1%	1 (0,253)	0 (1,000)	0 (1,000)	3 (0,003)
5%	3 (0,175)	1 (0,774)	1 (0,774)	7 (0,0004)
10%	3 (0,565)	2 (0,801)	4 (0,329)	7 (0,022)
Nº DE POSTOS	29	29	29	29

Tabela 2.5 - Número de séries rejeitadas pelo teste de Kendall na bacia do Rio Doce. Séries obtidas pelo critério b.

OBS.: Entre parênteses, a probabilidade de que o nº de séries rejeitadas seja maior ou igual ao nº observado.

Uma análise das Tabelas 2.4 e 2.5 mostra que os dois critérios de obtenção de séries parciais estudados foram capazes de fornecer séries aleatórias, pelo menos para uma taxa anual de até três pi-

cos por ano. Para a taxa de quatro picos por ano, os dois critérios apresentaram inconvenientes. As séries obtidas pelo critério a não garantiram uma taxa anual de picos igual a quatro em todos os postos da bacia, muito embora, nos postos onde isto foi possível, as séries obtidas apresentaram um grau de aleatoriedade aceitável. Já as séries obtidas pelo critério b foram suficientemente longas para fornecerem taxas de quatro picos por ano, só que, neste caso, apresentaram um grau de não-aleatoriedade acentuado. Vale ressaltar, no entanto, que não foram tomados maiores cuidados no sentido de melhorar a qualidade das séries obtidas. Por um lado, os cuidados propostos por NERC (1975) citados no item 2.3.2 poderiam fornecer séries com um grau menor de persistência. Isto acarretaria, entretanto, a necessidade de maiores informações sobre características geomorfológicas da bacia. Por outro lado, uma escolha mais cuidadosa do nível de truncamento a ser utilizado na obtenção das séries pelo critério a poderia eliminar o inconveniente das baixas taxas anuais de picos.

Um aspecto interessante deve ser ressaltado nas Tabelas 2.4 e 2.5: é de se esperar que o grau de não-aleatoriedade das séries parciais cresça à medida em que a taxa anual de picos é aumentada, uma vez que os picos da série passam a estar mais próximos entre si. Observa-se, entretanto, que o número de séries com $\lambda = 1$ rejeitadas pelo teste de aleatoriedade é maior que o número de séries com $\lambda = 2$ rejeitadas. Uma possível explicação para esta aparente contradição é que nas séries contendo em média apenas um pico por ano existem uns poucos anos contendo vários picos (dependentes entre si), fato que é "disfarçado" nas séries com dois ou mais picos por ano, mais extensas.

No restante deste trabalho as séries parciais serão obtidas sempre pelo critério dos máximos excessos, a menos que se diga o contrário. Esta escolha se prende a que, tendo sido di fícil uma escolha com bases mais objetivas, optou-se por adotar o critério que possui significado mais claro, do ponto de vista con ceitual.

CAPÍTULO III

SÉRIES DE VAZÕES EXTREMAS DA BACIA DO RIO DOCE

É extremamente recomendável a prática de se iniciar um trabalho de análise de frequência verificando os graus e tipos de variabilidade existentes nos dados disponíveis (Benjamin e Cornell, 1970). Na exposição dos métodos de análise de frequência de cheias, quer a partir de séries de máximos anuais, quer a partir de séries parciais, foram mencionadas diversas hipóteses associadas a cada método. No caso de séries de máximos anuais, as hipóteses citadas têm sido testadas para séries provenientes de bacias hidrográficas de várias regiões do país (Pinheiro, 1982; Damazio e Kelman, 1983), e os resultados mostram que elas são, em geral, aceitáveis. Já com relação às séries parciais, pouco tem sido feito neste sentido, talvez devido à sua pouca utilização, sendo interessante uma verificação prática das hipóteses levantadas no item 2.3.1.

3.1 - Características Gerais das Séries de Extremos

Com o objetivo de tornar familiares as peculiaridades relativas às séries parciais, são mostradas neste item algumas estatísticas das séries parciais dos postos fluviométricos da bacia do Rio Doce. Além disso, para permitir uma comparação entre as características destas séries e as das de máximos anuais correspondentes, são também apresentadas as estatísticas destas últimas.

Foram estudadas quatro diferentes séries de extre-

mos em cada registro. A série de máximos anuais (SMA) e as séries parciais com taxas anuais de picos, λ , iguais a 1, 2 e 3 (SP1, SP2 e SP3, respectivamente). As séries parciais foram obtidas pelo critério dos máximos excessos (critério a, item 2.3.2), fixando-se como nível de truncamento a metade da média dos máximos anuais em cada registro e selecionando-se respectivamente os N, 2N e 3N maiores picos de cada uma, sendo N o comprimento do registro disponível. Em apenas um dos 29 postos considerados não se conseguiu, com o nível de truncamento adotado, uma série com uma média de três picos por ano.

3.1.1-Estatísticas Locais e Regionais

Uma primeira idéia sobre as possíveis distribuições de probabilidades das populações pode ser formada a partir do cálculo dos momentos amostrais. Foram calculados para as 4 séries de cada posto a média e os coeficientes adimensionais de variação e assimetria. Os resultados obtidos aparecem na tabela 3.1, onde se pode observar que, de maneira geral, as séries parciais apresentam assimetria e curtose maiores que as séries de máximos anuais, enquanto que para os coeficientes de variação ocorre o contrário. Para as séries parciais, os coeficientes de variação e de assimetria crescem à medida que se aumenta a taxa anual de picos.

Uma forma de incorporar informações contidas em vários registros de uma mesma região é a determinação de coeficientes adimensionais regionais de variação, assimetria e curtose. Naturalmente, estes coeficientes podem fornecer algum tipo de informação apenas se a região puder ser considerada homogênea do ponto de vista hidrológico. NERC (1975) propõe as seguintes expressões

POSTO: CÓDIGO DNAEE	ÁREA DE DRENAGEM Km ²	SMA			SP1			SP2			SP3		
		MÉDIA m ³ /s	CV	ASS	MÉDIA m ³ /s	CV	ASS	MÉDIA m ³ /s	CV	ASS	MÉDIA m ³ /s	CV	ASS
56010000	540	64	0,32	0,84	77	0,18	1,64	65	0,25	1,24	57	0,30	1,23
56028000	1.415	182	0,70	1,99	230	0,51	1,98	180	0,54	2,81	151	0,50	3,12
56050000	333	39	0,49	0,34	53	0,19	0,47	43	0,28	0,88	37	0,35	1,09
56075000	4.171	403	0,44	1,60	460	0,30	2,24	372	0,36	2,35	326	0,39	2,49
56090000	331	15	0,21	- 0,60	16	0,12	0,40	14	0,19	0,68	13	0,23	0,83
56110000	6.197	538	0,50	1,37	621	0,36	1,93	493	0,41	2,34	429	0,44	2,60
56145000	213	51	0,48	0,78	58	0,34	1,37	44	0,44	1,72	37	0,50	1,98
56148000	303	128	0,52	0,37	171	0,28	0,60	127	0,44	1,03	105	0,52	1,39
56240000	843	172	0,48	0,86	197	0,32	1,34	156	0,39	1,78	136	0,42	2,11
56335000	1.301	244	0,44	0,99	274	0,31	1,52	220	0,36	1,93	194	0,39	2,24
56337000	523	79	0,30	0,97	86	0,21	1,99	74	0,24	2,03	66	0,28	1,79
56485000	1.306	133	0,57	1,76	151	0,51	1,47	116	0,56	2,35	98	0,59	2,81
56500000	268	32	0,52	1,06	39	0,32	1,48	30	0,42	1,71	25	0,49	1,86
56565000	307	14	0,15	- 1,22	15	0,08	1,19	13	0,13	0,50	12	0,19	0,35
56670000	527	194	0,41	- 0,61	225	0,19	0,30	173	0,36	0,54	144	0,46	0,92
56720000	24.245	1.565	0,41	0,86	1.784	0,30	1,23	1.434	0,36	1,66	1.229	0,42	1,80
56925000	160	16	0,59	2,08	18	0,48	2,55	14	0,51	3,19	12	0,53	3,58
56940000	3.181	156	0,53	0,93	186	0,37	1,12	147	0,42	1,86	126	0,47	2,12
56948000	61.610	2.459	0,28	0,54	2.612	0,20	1,39	2.170	0,27	1,33	1.883	0,34	1,30
56967000	1.521	249	0,71	1,42	302	0,53	1,38	218	0,64	2,11	183	0,68	2,64
56974000	2.092	244	0,57	1,44	273	0,45	1,57	212	0,50	2,33	182	0,53	2,75
56986000	313	55	0,46	1,24	62	0,33	1,86	49	0,38	2,22	43	0,42	2,41
56988500	1.345	161	0,45	1,49	177	0,36	1,93	139	0,42	2,26	119	0,47	2,48
56989500	3.375	322	0,65	2,49	386	0,58	1,94	288	0,65	2,77	243	0,68	3,30
56990000	8.454	509	0,40	0,64	544	0,30	0,91	435	0,37	1,54	382	0,40	1,91
56992000	2.162	174	0,60	1,11	207	0,44	1,13	153	0,55	1,83	-	-	-
56993600	1.100	32	0,31	- 0,19	36	0,19	0,36	30	0,23	1,12	28	0,25	1,47
56994000	774	56	0,54	1,66	63	0,43	1,80	51	0,44	2,60	45	0,45	2,97
56994500	76.615	4.007	0,28	0,11	4.299	0,18	1,20	3.551	0,27	1,03	3.092	0,33	1,17

Tabela 3.1 - Estatísticas das Séries de Extremos da Bacia do Rio Doce.

para o cálculo destes coeficientes regionais, para um região con-
tendo N séries de vazões extremas:

$$\overline{CV} = \left[\frac{\sum_{j=1}^N (n_j - 1) CV_j^2}{\sum_{j=1}^N (n_j - 1)} \right]^{1/2} \quad (3.1)$$

$$\overline{ASS} = \frac{\sum_{j=1}^N (n_j - 1) ASS_j}{\sum_{j=1}^N (n_j - 1)} \quad (3.2)$$

$$\overline{CUR} = \frac{\sum_{j=1}^N (n_j - 1) CUR_j}{\sum_{j=1}^N (n_j - 1)} \quad (3.3)$$

onde:

- CV_j , ASS_j e CUR_j são, respectivamente, os coeficientes de
variação, assimetria e curtose da j-ésima série que contém n_j ob-
servações.

- \overline{CV} , \overline{ASS} e \overline{CUR} são os coeficientes regionais de va-
riação, assimetria e curtose, respectivamente.

Estes coeficientes foram calculados para as 29 sê-
ries da bacia do Rio Doce, para cada um dos quatro tipos de sê-
ries de extremos, estando os resultados apresentados na tabela
3.2.

COEFICIENTES	TAXA ANUAL DE PICOS (PICOS/ANO)			SÉRIES DE MÁXIMOS ANUAIS
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	
Variação	0,318	0,374	0,457	0,484
Assimetria	2,015	2,453	2,771	1,554
Curtose	8,798	11,408	13,515	6,999
Nº de Picos	719	1438	2157	719

Tabela 3.2 - Estatísticas Regionais das Séries de
Extremos da bacia do Rio Doce

Pode-se observar que as séries de máximos anuais desta bacia apresentam um coeficiente de variação regional maior que os coeficientes dos três tipos de séries parciais correspondentes, para as quais as menores taxas de picos correspondem aos menores coeficientes de variação. Já os coeficientes de assimetria e curtose das séries parciais são superiores aos da série de máximos anuais, sendo seus valores crescentes com a taxa anual de picos.

Na Figura 3.1 são apresentadas curvas que representam a relação teórica entre os coeficientes de assimetria e de curtose de algumas distribuições de probabilidades: gama, log-normal, Gumbel generalizada, Gumbel e exponencial. Neste gráfico estão também apresentados os pontos representativos das relações empíricas entre estes dois coeficientes, obtidas para a bacia do rio Doce. Nota-se que, para as três séries parciais, os pontos amostrais se aproximam bastante da distribuição gama, sendo o ponto correspondente à taxa de um pico por ano muito próximo da distri-

buição exponencial. Este resultado fica interessante se for analisado da seguinte forma: à medida que o nível de truncamento da série parcial cresce (λ diminui), a distribuição de probabilidades dos picos se aproxima da exponencial. Isto lembra um resultado teórico obtido por Shane e Lynn (1964), citado anteriormente. Entretanto, não se pode esquecer que os resultados apresentados aqui são específicos da bacia do rio Doce e não permitem, a priori, nenhum tipo de conclusão generalizada. Quanto à série de máximos anuais, o ponto correspondente a sua relação assimetria/curtose se encontra entre as curvas das distribuições gama e log-normal.

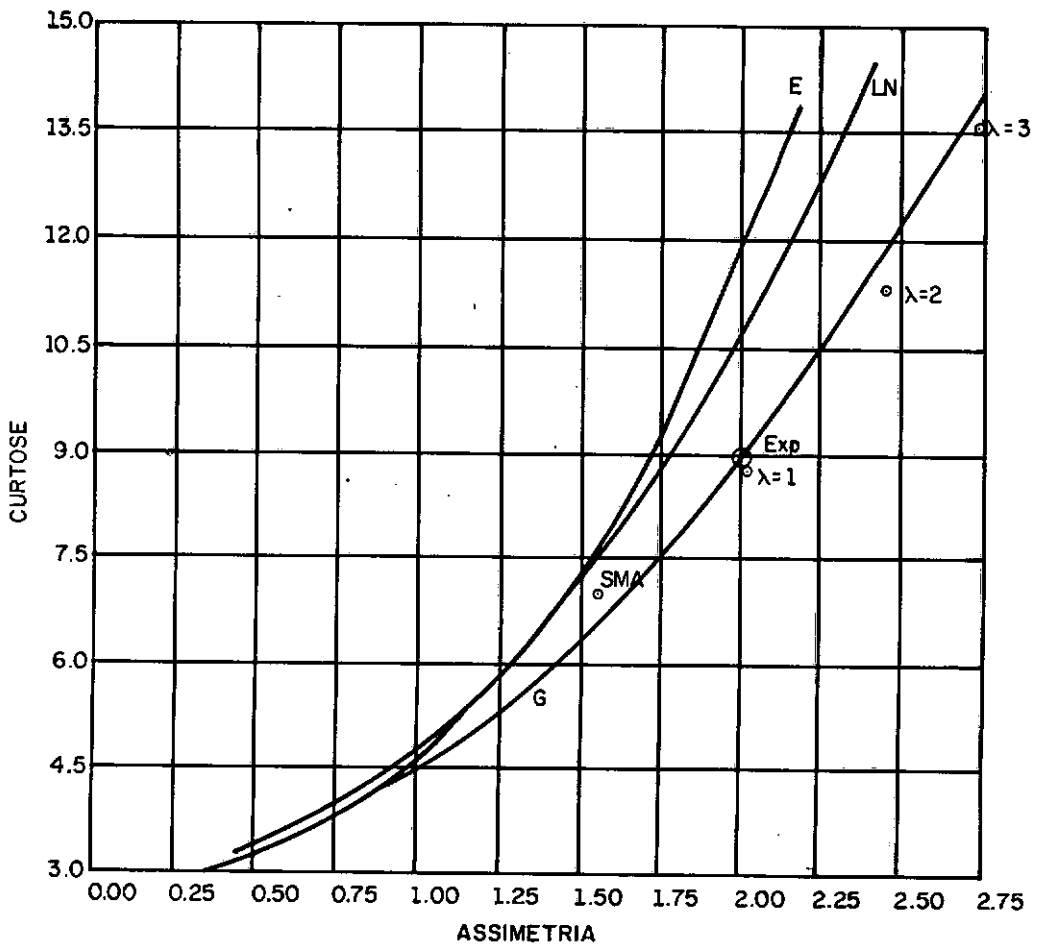


FIGURA 3.1 — COEFICIENTES REGIONAIS PARA A BACIA DO RIO DOCE

G — DISTRIBUIÇÃO GAMA

LN — DISTRIBUIÇÃO LOG - NORMAL

E — DISTRIBUIÇÃO DE EXTREMOS

Exp — DISTRIBUIÇÃO EXPONENCIAL

SMA — SÉRIE DE MÁXIMOS ANUAIS

$\lambda=1, \lambda=2, \lambda=3$ — SÉRIES PARCIAIS

3.1.2-Modelos de Regressão Linear

Os valores médios das séries de vazões extremas de uma bacia hidrográfica variam devido não só à própria aleatoriedade do fenômeno como também às diferentes características hidrológicas das bacias de drenagem dos postos fluviométricos. Por exemplo, é razoável supor que a ordem de grandeza da atividade hidrológica cresce à medida que cresce a área de drenagem. Usualmente, a quantificação deste efeito das características da bacia hidrológica sobre as vazões médias é feita através de modelos de regressão linear múltipla, com expressões do tipo:

$$\bar{q} = c A^K \cdot \prod_{i=1}^n \theta_i^{K_i} \quad (3.4)$$

onde \bar{q} é a média das vazões extremas, c e K coeficientes empíricos, A é a área de drenagem da bacia e θ_i são termos representativos de outras características físicas da bacia hidrográfica. Sem dúvida, a área de drenagem, A , é o fator mais importante da equação 3.4, sendo comum ignorarem-se os outros termos, fazendo $K_i = 0$; $i=1, 2, \dots, N$ (NERC, 1975). Assim, tem-se a expressão mais simples:

$$\bar{q} = c A^K \quad (3.5)$$

na qual os coeficientes c e K são geralmente ajustados pelo método dos mínimos quadrados.

A equação 3.5 foi ajustada aos 29 postos da bacia do rio Doce, sendo consideradas as quatro séries de extremos: SMA, SP1, SP2 e SP3. Os coeficientes c e K obtidos para cada série es

tão listados na tabela 3.3, onde aparecem também os coeficientes de determinação e os erros fatoriais (NERC, 1975) das equações que quantificam a qualidade dos ajustes obtidos. Pode-se verificar que as séries parciais mais extensas (2 e 3 picos por ano) forneceram ajustes melhores que as séries de máximos anuais, sendo que, de maneira geral, os 4 ajustes podem ser considerados satisfatórios. Pode-se concluir, portanto, que para a bacia do rio Doce existe uma boa relação empírica entre as médias das séries de extremos e as áreas de drenagem.

Para as séries parciais, existe ainda uma outra variável, cujo valor pode ser associado, de forma semelhante, às áreas de drenagem: o nível de truncamento da série. Foram tomadas as séries parciais dos 29 postos da bacia do rio Doce com as três taxas anuais de picos em estudo e, para cada tipo de série, ajustada uma equação do tipo:

$$q_0 = c_0 A^{K_0} \quad (3.6)$$

onde q_0 é o nível de truncamento de uma série parcial de N anos contendo $\lambda.N$ picos, A é a área de drenagem da bacia e c_0 e K_0 são coeficientes empíricos. Esta relação é interessante, pois associa a área da bacia hidrográfica um possível limite inferior para a distribuição de probabilidades dos picos. Os resultados obtidos são apresentados na tabela 3.4. Uma comparação destes resultados com os da tabela anterior mostra que os ajustes obtidos para equação 3.6 foram superiores aos da equação 3.5, ou seja: para esta região existe, aparentemente, uma relação estreita ligando a área de drenagem das bacias ao limite inferior de suas séries parciais. Note-se que esta constatação poderia ser aproveitada sob um outro

COEFICIENTES	TAXA ANUAL DE PICOS (PICOS / ANO)			SÉRIE DE MÁXIMOS ANUAIS
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	
C	0,492	0,384	0,328	0,403
K	0,811	0,817	0,822	0,810
R ²	0,84	0,86	0,87	0,85
EF%	80	73	69	78

Tabela 3.3 - Regressão: Média dos Picos x Áreas de Drenagem $\bar{q} = \bar{c}A^K$.

COEFICIENTES	TAXA ANUAL DE PICOS (PICOS / ANO)		
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$
C ₀	0,341	0,241	0,192
K ₀	0,817	0,831	0,839
R ²	0,86	0,89	0,90
EF%	71	61	57

Tabela 3.4 - Regressão: Nível de Truncamento x Área de Drenagem $q_0 = c_0 A^{K_0}$

Observação: R² = Coeficiente de Determinação
 EF = Erro Fatorial
 q₀ = Nível de Truncamento (m³/s)
 \bar{q} = Média dos Picos (m³/s)
 A = Área de Drenagem (km²)

aspecto: dada a área de drenagem associada a um posto fluviométrico, é possível, a partir de fórmulas empíricas e conhecendo-se apenas a área de drenagem da bacia, definir uma distribuição exponencial de dois parâmetros para as séries parciais deste posto. Para este ajuste, os dois parâmetros são o nível mínimo e a média da série, obtidos a partir das equações 3.5 e 3.6.

3.2 - Hipóteses dos Modelos de Séries Parciais

3.2.1-A Hipótese Poisson

A hipótese de que, nas séries parciais, o número de picos ocorridos em um ano tem uma distribuição Poisson tem sido adotada em diversos estudos teóricos (Shane e Lynn, 1964; Zelenhasic, 1970). Numa tentativa de se encontrar alguma evidência empírica acerca desta hipótese, foram obtidas séries parciais com diferentes níveis de truncamento para os postos da bacia do rio Doce, sendo calculada, para cada série, a relação amostral média/variância da variável "número de picos por ano". Como se sabe, para a distribuição Poisson, esta relação é igual à unidade. Os resultados obtidos aparecem condensados na Figura 3.2, num gráfico que mostra o número médio de picos por ano nas abscissas e a relação média/variância nas ordenadas. Pode-se ver que apenas para as taxas anuais de picos mais baixas (maiores níveis de truncamento) é que se verificaram, para as séries estudadas, relações média/variância mais próximas da unidade. Não existe, portanto, nenhuma evidência empírica de que o número de picos por ano tenha distribuição Poisson. Entretanto, como já foi dito anteriormente, este resultado não constitui um problema para a determinação de vazões com tempos de recorrência elevados pelos métodos de séries

parciais, uma vez que, neste caso, é suficiente que se conheça o valor esperado do número de picos por ano.

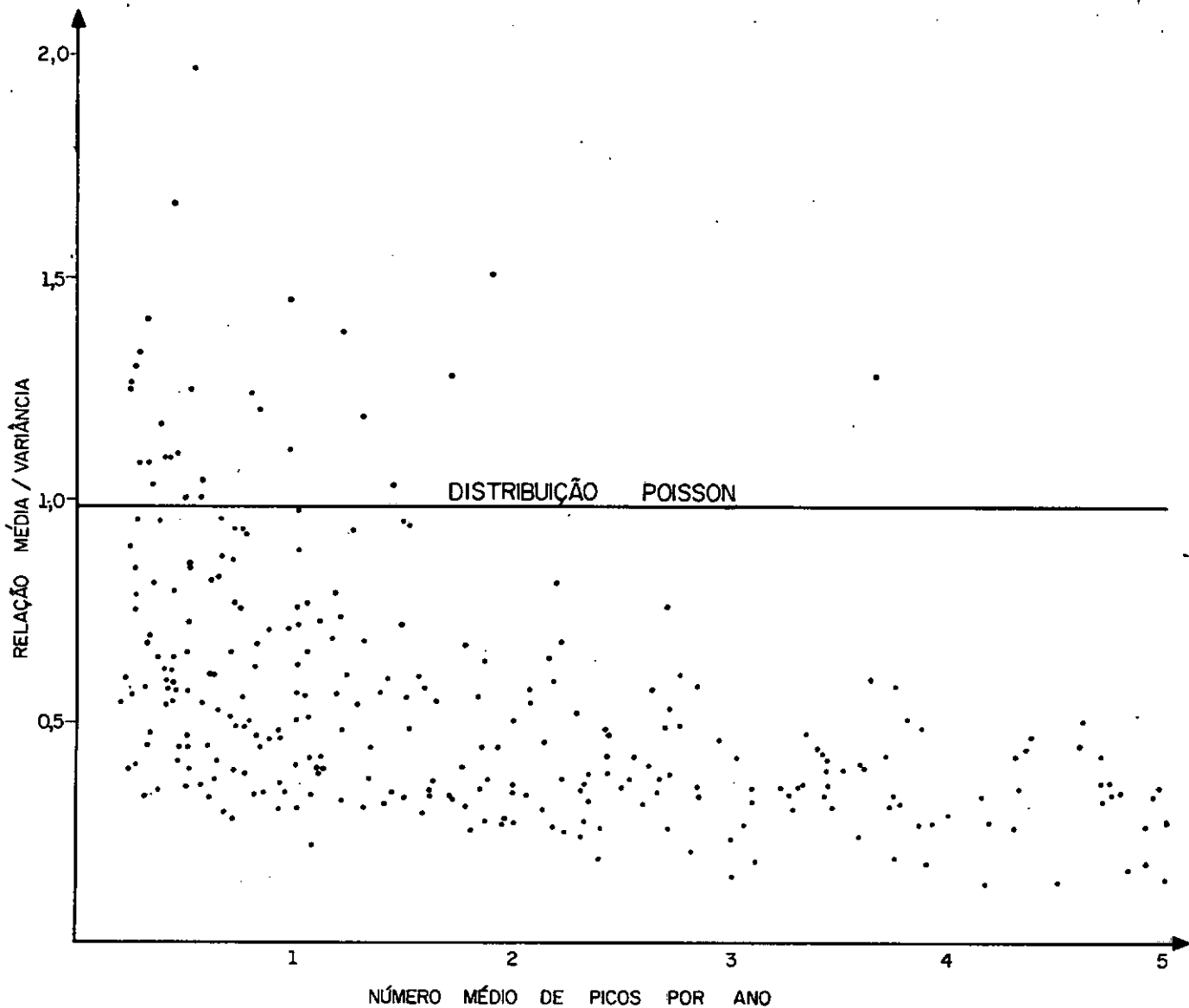


FIGURA 3.2 — VARIÁVEL "NÚMERO DE PICOS A CADA ANO" — RELAÇÃO ENTRE A MÉDIA E A VARIÂNCIA — BACIA DO RIO DOCE.

3.2.2-A Hipótese Exponencial

A hipótese de que os picos de cheias são provenientes de uma população cuja distribuição de probabilidades apresenta, na cauda direita, um comportamento idêntico ao da distribui-

ção exponencial tem sido freqüentemente adotada na literatura. Recentemente, Kavvas (1982) propôs um modelo de séries parciais relativamente complexo. Este modelo assume, entre outras, a hipótese de que o número médio de picos acima de um nível de truncamento, ocorridos ao longo de um dado intervalo de tempo, decresce exponencialmente à medida que este nível é elevado. Esta hipótese pode ser expressa analiticamente por:

$$\lambda = a \exp (- b q_0) \quad (3.7)$$

sendo λ o número médio de picos, q_0 o nível de truncamento da série e a e b coeficientes empíricos a serem determinados. As considerações a seguir mostram algumas implicações desta hipótese.

Considere-se uma série parcial cujo nível de truncamento seja baixo o suficiente para que ela contenha todos os picos ocorridos. Seja M a variável aleatória "número de picos ocorridos num ano qualquer", e $N(x)$ o número destes picos cuja magnitude seja superior a x . A distribuição de probabilidades da variável "magnitude dos picos" é $F(x)$ e a probabilidade de que um pico qualquer seja maior que x , p_x , é dada por:

$$p_x = 1 - F(x) \quad (3.8)$$

Se em um determinado ano ocorrem m picos, e estes picos são independentes, o valor esperado do número destes picos maiores que x pode ser dado por:

$$E \{ N(x) \mid M = m \} = p_x m \quad (3.9)$$

Considerando-se agora a variável m ,

$$E\{N(x)\} = \sum_{m=0}^{\infty} p_x \cdot m \cdot P\{M = m\} = p_x \sum_{m=1}^{\infty} m \cdot P\{M = m\} \quad (3.10)$$

ou seja:

$$E\{N(x)\} = p_x \cdot E\{M\} \quad (3.11)$$

Se a hipótese descrita pela equação 3.7 for aceita, é possível escrever:

$$\hat{E}\{N(x)\} = \hat{\lambda} = a \cdot \exp(-b x) \quad (3.12)$$

se \bar{M} for uma estimativa do valor esperado de M , λ , pode-se, a partir das equações 3.11 e 3.12, escrever

$$\hat{p}_x \cdot \bar{M} = a \cdot \exp(-b x) \quad (3.13)$$

e obter uma estimativa de p_x dada por:

$$\hat{p}_x = \frac{a}{\bar{M}} \cdot \exp(-b x) \quad (3.14)$$

Substituindo-se este resultado na equação 3.8, obtém-se a seguinte expressão para a distribuição de probabilidades da variável "magnitude dos picos":

$$\hat{F}(x) = 1 - \frac{a}{M} \exp(-b x) \quad (3.15)$$

Através de simples transformação algébrica desta equação, chega-se à expressão

$$F(x) = 1 - \exp \left\{ -b \left[x + \frac{1}{b} \ln \left(\frac{a}{M} \right) \right] \right\} \quad (3.16)$$

que é idêntica à expressão da distribuição exponencial com a origem deslocada,

$$F(x) = 1 - \exp \left[-\beta(x - x_0) \right] \quad (3.17)$$

sendo:

$$\left. \begin{aligned} \beta &= b \\ x_0 &= -\frac{1}{b} \ln \left(\frac{a}{M} \right) \end{aligned} \right\} \quad (3.18)$$

A hipótese de que o número de picos maiores que um certo nível decresça exponencialmente à medida que este nível aumenta, é, então, equivalente à hipótese de que a magnitude dos picos tenha uma distribuição exponencial. As equações 3.18 fornecem um expediente alternativo de se estimarem os parâmetros de tal distribuição. Os valores de a e b podem ser obtidos pelo método dos mínimos quadrados, e o valor de M calculado diretamente da amostra.

Os coeficientes a e b foram determinados para as séries parciais da bacia do rio Doce ajustando-se a equação 3.7 pelo método dos mínimos quadrados. Os coeficientes de determinação das regressões obtidas foram altos, sendo que em apenas dois dos 29 postos foram encontrados valores inferiores a 0,85. Este re-

sultado pode ser considerado uma evidência empírica razoável de que a hipótese descrita na equação 3.7 é viável.

Entretanto, para conclusões mais abalizadas, seria interessante que a hipótese exponencial fosse verificada através de um teste de hipóteses. Um teste para a hipótese de que uma determinada amostra provém de uma população exponencialmente distribuída foi proposto por Gnedenko et al. (1969) tendo já sido utilizado em aplicações hidrológicas por Kottegoda (1979), no estudo das caudas das distribuições de vazões diárias. Este teste é descrito a seguir.

Hipótese nula, H_0 : A variável aleatória X é exponencialmente distribuída, ou seja, sua distribuição de probabilidades é dada pela expressão 3.17 sendo β constante.

Esta hipótese é testada contra duas hipóteses alternativas: a variável X tem uma distribuição cuja cauda direita é (1) menos densa (cauda leve) ou (2) mais densa (cauda pesada) que a da distribuição exponencial. Estas duas hipóteses alternativas seriam caracterizadas por ser β uma função de x (1) monotonamente crescente ou (2) monotonamente decrescente.

O teste é aplicado da seguinte forma:

1. São escolhidos de uma amostra da população X os valores superiores a um certo nível x_0 , obtendo-se assim uma sub-amostra de n elementos. Kottegoda (1979) sugere que se defina x_0 de forma que n seja correspondente a menos que 5% dos elementos da amostra inicial.

2. A sub-amostra \tilde{e} ordenada em um vetor $(x(1), x(2), \dots, x(n))$ de forma que $x(1) \leq x(2) \leq \dots \leq x(n)$.

3. Define-se uma nova variável $S(i)$, da seguinte forma:

$$S(i) = (n-i+1) \left[x(i) - x(i-1) \right] \quad (3.19)$$

para $i = 1, 2, \dots, n$, com $x(0) = x_0$.

4. A estatística do teste \tilde{e} obtida dividindo-se o vetor $x(i)$ em dois grupos, $(x(1), x(2), \dots, x(r))$ e $(x(r+1), x(r+2), \dots, x(n))$ e calculando:

$$Q(r, n-r) = \frac{\sum_{i=1}^r [S(i)/r]}{\sum_{i=r+1}^n [S(i)/(n-r)]} \quad (3.20)$$

Se a hipótese nula for verdadeira, esta estatística tem uma distribuição de probabilidades $F [V1, V2]$ com graus de liberdade $V1 = 2r$ e $V2 = 2(n-r)$. A hipótese nula não é rejeitada, para um certo nível de significância α , se as duas desigualdades seguintes forem verificadas simultaneamente:

$$Q(r, n-r) \leq F_{(1-\alpha/2)} \left[2r, 2(n-r) \right] \quad (3.21)$$

$$\left[Q(r, n-r) \right]^{-1} \leq F_{(1-\alpha/2)} \left[2(n-r), 2r \right] \quad (3.22)$$

Os valores à direita das duas inequações anteriores podem ser obtidos em uma tabela de distribuição F (Mood et al.; 1974; Gnedenko et al., 1969). Caso a desigualdade 3.21 seja violada, rejeita-se a hipótese nula em favor da alternativa de que β é uma função crescente de x , e a variável X tem cauda leve. Caso a segunda desigualdade, 3.22, seja violada, rejeita-se a hipótese nula em favor da alternativa de que β é uma função decrescente de x , ou seja, a distribuição de X possui cauda pesada.

Aplicou-se o teste de Gnedenko às séries parciais da bacia do rio Doce. Para que fossem considerados apenas os valores da cauda da distribuição, foram tomadas amostras contendo apenas os 20 maiores picos de cada série. Com o objetivo de se obterem estatísticas mais sensíveis ao comportamento da cauda, considerou-se $r=18$ (Bryson, 1974) em todos os casos.

A tabela 3.5 mostra os postos fluviométricos em cujas séries a hipótese nula foi rejeitada, em favor da de cauda leve, e os níveis de significância da decisão. Se os picos realmente fossem exponencialmente distribuídos em todos os 29 postos da bacia, haveria uma probabilidade de 33% de que a hipótese nula fosse rejeitada em 4 postos a um nível de significância de 10% (ou uma probabilidade de 25% de que fosse rejeitada em um posto). Logo, não há evidências de que a variável aleatória "magnitude dos picos" tenha distribuição com cauda leve nas séries parciais da bacia do rio Doce. Quanto à cauda pesada, em apenas um posto (código DNAEE 56075000) a hipótese nula foi rejeitada em favor desta alternativa, para um nível de significância de 10%.

CÓDIGO DNAEE DO POSTO	NÍVEL DE SIGNIFICÂNCIA
56148000	10%
56948000	1%
56990000	10%
56994500	10%

Tabela 3.5 - Postos cujos picos possuem distribuição com cauda leve, segundo o teste de Gnedenko

Pode-se concluir, portanto, que, para as séries parciais da bacia do rio Doce, não há evidências empíricas de que os picos tenham uma distribuição de probabilidades com cauda diferente da distribuição exponencial. Esta constatação é importante, na medida que fecha o leque das possíveis distribuições a serem utilizadas na análise de vazões extremas desta bacia.

3.3 - Regionalização de Informações

A análise de frequência de cheias num local onde o registro de vazões disponível fosse extremamente longo poderia ser efetuada tendo em vista apenas os dados contidos neste registro que poderiam, neste caso, fornecer todas as informações necessárias à escolha de uma distribuição de frequência de extremos adequada. Entretanto, na maioria dos casos, o registro de vazões disponível não tem grande extensão e, em geral, sempre é vantajoso lançar mão de informações contidas no conjunto de registros de outros postos da vizinhança. Naturalmente, para que estas informações regionais sejam eficientes, é preciso que todos estes postos se situem numa região homogênea do ponto de vista hidrológico, sem

o que poder-se-iam introduzir tendências nos resultados. Além disso, se houver uma grande dependência entre estes registros de vazões, as informações fornecidas pelas considerações regionais podem acrescentar pouco ou nada às informações locais já disponíveis. Assim, o que se quer é um conjunto de postos fluviométricos situados em locais hidrologicamente semelhantes sem, contudo, que seus registros apresentem grande dependência.

3.3.1-Distribuições Regionais de Frequência

Uma forma de se considerar as informações regionais é o estabelecimento de estatísticas regionais, conforme feito no item 3.1.1 para a bacia do rio Doce (coeficientes de variação, as simetria e curtose), a partir dos registros de extremos disponíveis. A média da série de extremos num certo local da região pode ser obtida:

- 1) aplicando-se as equações 3.4 ou 3.5, conforme o caso, se não se dispõe de registro algum no local, ou
- 2) a partir do registro de extremos existente, mesmo de pequena extensão, estimando o valor esperado como sendo a média amostral.

Neste último caso, NERC (1975) sugere que se estime a média dos máximos anuais a partir das séries parciais, sempre que os registros disponíveis forem de pequena extensão (menos de dez anos). Estimada a média desta forma, pode-se ajustar, pelo método dos momentos, uma distribuição de probabilidades adequada para a série

de vazões extremas fazendo:

$$\hat{C\bar{V}} = \overline{C\bar{V}} \quad (\text{eq. 3.1})$$

$$\hat{A\bar{S}S} = \overline{A\bar{S}S} \quad (\text{eq. 3.2})$$

$$\hat{C\bar{U}R} = \overline{C\bar{U}R} \quad (\text{eq. 3.3})$$

A escolha da distribuição mais adequada pode ser feita com base em considerações conceituais ou empíricas. No caso de séries parciais, por exemplo, a distribuição exponencial se aplica mediante algumas hipótese já discutidas no item II.3.1. Aceitas estas mesmas hipóteses, teoricamente a distribuição dos máximos anuais deveria ser Gumbel. De forma empírica, por outro lado, NERC (1975) sugere que seja investigada a distribuição de probabilidades que melhor se ajusta às séries de extremos da região, sendo esta distribuição considerada a mais adequada também para os locais em que os registros são curtos ou inexistentes.

O modelo de regressão linear pode ser utilizado para estimar os parâmetros de uma distribuição exponencial regional para as séries parciais. As expressões 3.5 e 3.6 fornecem a média e o nível de truncamento das séries parciais, em função da área de drenagem da bacia hidrográfica. Estes dois valores podem ser usados como estimadores do limite inferior, x_0 , e da média $\beta + x_0$, da distribuição exponencial, que será dada por

$$F(x) = 1 - e^{-\frac{1}{\beta}(x-x_0)} \quad (3.23)$$

As tabelas 3.5 e 3.6 fornecem os coeficientes da regressão para a determinação dos dois parâmetros, β e x_0 , para os postos da bacia

do rio Doce. Este procedimento é interessante porque permite a estimação de vazões com diferentes tempos de recorrência, tomando como dado apenas a área da bacia.

Uma abordagem alternativa para esta transferência de informações a nível regional é a determinação de uma distribuição empírica de freqüência de extremos para a região. Naturalmente, como a ordem de grandeza das vazões afluentes a um posto varia de acordo com a área de drenagem associada, é preciso se considerar vazões reduzidas ou padronizadas, de modo que se possa tratar todos os registros de forma conjunta. Esta padronização de vazões é feita geralmente

- 1) dividindo-se as vazões de cada posto pela área de drenagem do mesmo (vazões específicas) ou
- 2) dividindo-se as vazões de cada série pela sua média (vazões adimensionais).

Estas vazões padronizadas, $Q_e = Q/\bar{A}$ ou $Q_a = Q/Q_{med}$, de todos os postos são ordenadas e marcadas em um papel de probabilidades, permitindo assim o estabelecimento da distribuição empírica das vazões extremas para a região, através da união dos pontos e, se necessário, da extrapolação da curva resultante.

Foram traçados gráficos das distribuições regionais de freqüência para os quatro tipos de séries de extremos em estudo na bacia do rio Doce: SMA, SP1, SP2 e SP3. O método utilizado foi o proposto por NERC (1975), sumariamente descrito a seguir:

1. Cada série j de vazões extremas é tornada adimensional, dividindo-se seus elementos pelo seu valor médio:

$$x_i = \frac{q_i}{\bar{q}_j} ; i = 1, 2, \dots, n_j \quad (3.24)$$

sendo n_j o número de elementos da j -ésima série e $j=1, 2, \dots, 29$.

2. Os elementos de cada uma destas séries adimensionais são ordenados em um conjunto $\{x_j(1), x_j(2), \dots, x_j(n_j)\}$ de forma que $x_j(1) < x_j(2) < \dots < x_j(n_j)$.
3. Cada valor $x_j(i)$ é associado à probabilidade $p_j(i)$ de ocorrência de valores menores que ele na população, através de um estimador. Nesta análise, assumiu-se que as populações tivessem distribuição Gumbel. Neste caso, uma estimativa não tendenciosa de $p_j(i)$ é dada pela fórmula de Gringorten:

$$p_j(i) = P \{X \leq x_j(i)\} = \frac{i - 0,44}{n + 0,12} \quad (3.25)$$

4. A cada valor destes é associado um ponto num gráfico cartesiano onde nas ordenadas estão os valores adimensionais $x_j(i)$ e nas abscissas as probabilidades $p_j(i)$. Para que a distribuição Gumbel possa ser traçada como uma reta neste gráfico, os valores das abscissas devem ser:

$$y(i) = -\ln \{ -\ln [p(i)] \} \quad (3.26)$$

5. Como o número de elementos n_j é variável de série para série, o gráfico resultante apresenta uma nuvem de pontos que pode

ser mais ou menos dispersa. Para se evitar o acúmulo de pontos neste gráfico, dividem-se as abscissas em intervalos $(-1,5; -1,0)$, $(-1,0; -0,5)$, $(-0,5; -0,0)$, ... etc. Para cada intervalo marca-se apenas um ponto médio, com coordenadas iguais às médias das coordenadas dos pontos contidos neste intervalo. Deve-se ressaltar que a escolha da distribuição Gumbel, para a escala do gráfico ou da posição de plotagem, não implica na aceitação desta distribuição para a modelagem das vazões extremas.

As figuras 3.3 a 3.6 apresentam os quatro gráficos representativos das distribuições empíricas de probabilidades das quatro séries de extremos da bacia do rio Doce. Na escala horizontal, além de serem marcados os valores relativos à variável reduzida Gumbel, foram indicados os pontos correspondentes a alguns tempos de recorrência. No caso das séries parciais, estes pontos já levam em conta o valor esperado do número de picos por ano, ou seja: $T=100$ equivale ao valor de SP3 com probabilidade $1/300$ de ser excedido.

O problema da aplicação destas curvas na estimação de vazões com grande tempos de recorrência é que a extrapolação das curvas obtidas é dificultada pela má definição dos pontos à direita, justamente a região mais importante.

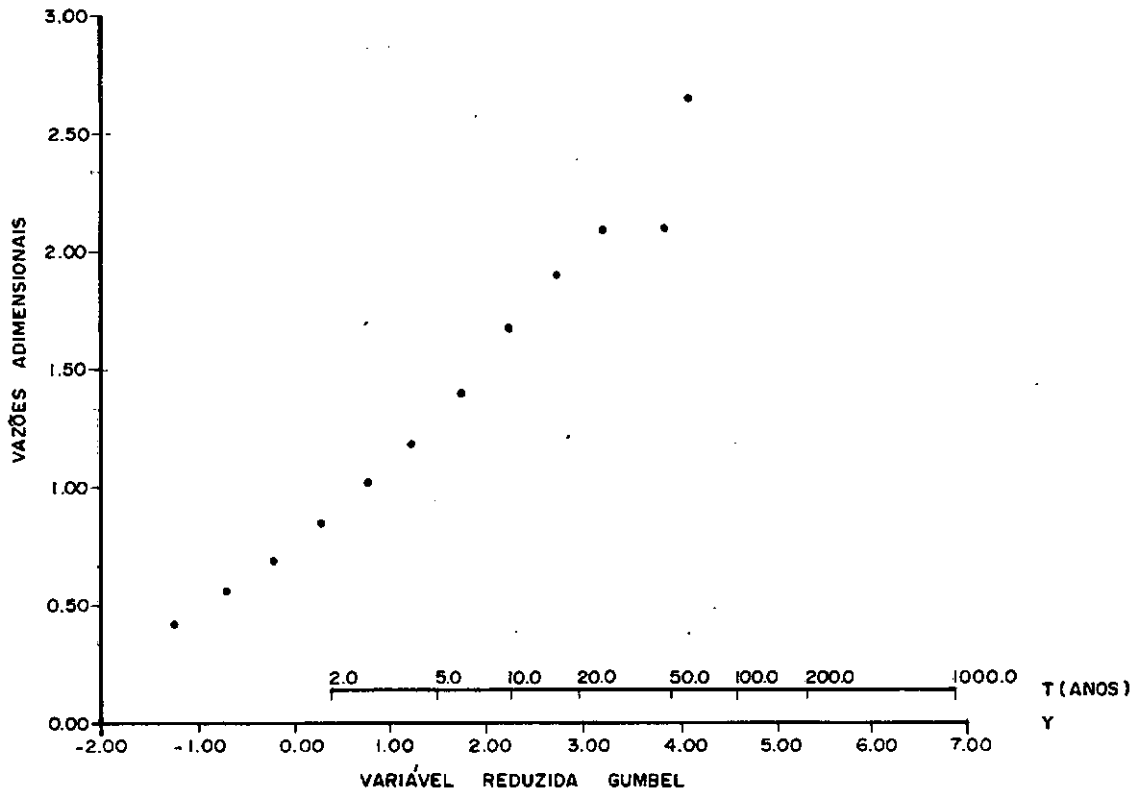


FIGURA 3.3 — DISTRIBUIÇÃO EMPÍRICA REGIONAL DE MÁXIMOS ANUAIS — BACIA DO RIO DOCE.

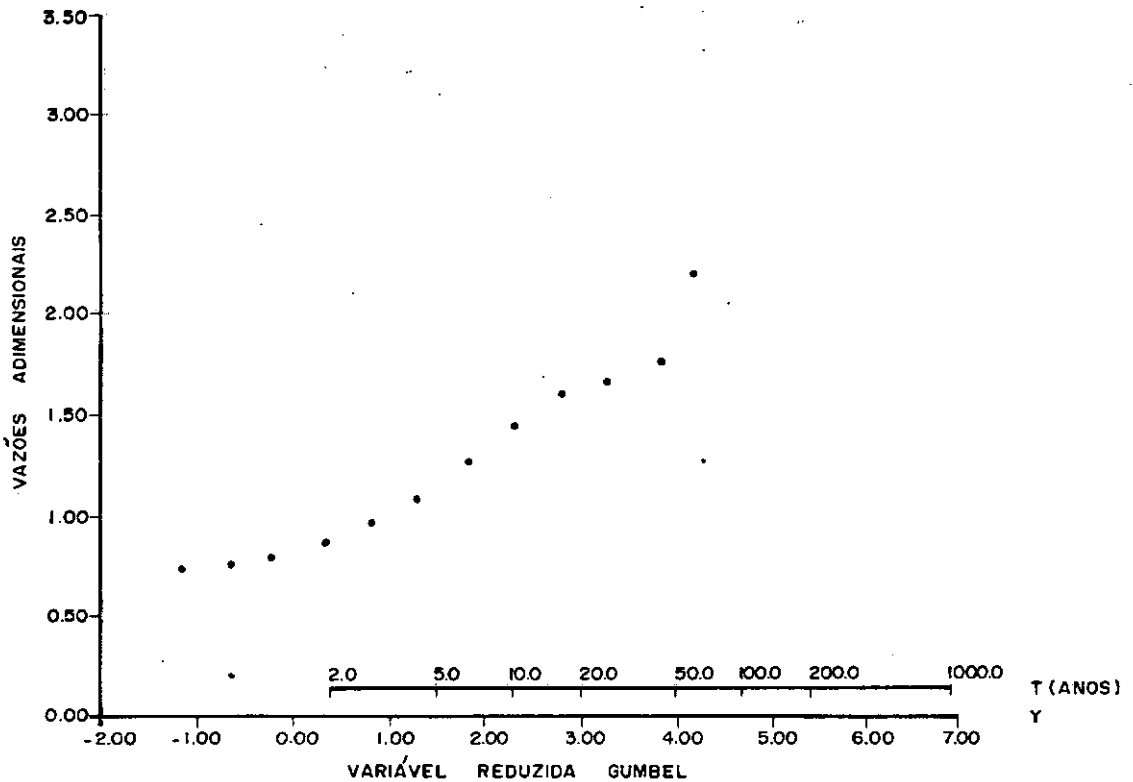


FIGURA 3.4 — DISTRIBUIÇÃO EMPÍRICA REGIONAL DAS SÉRIES PARCIAIS — SP 1 — BACIA DO RIO DOCE.

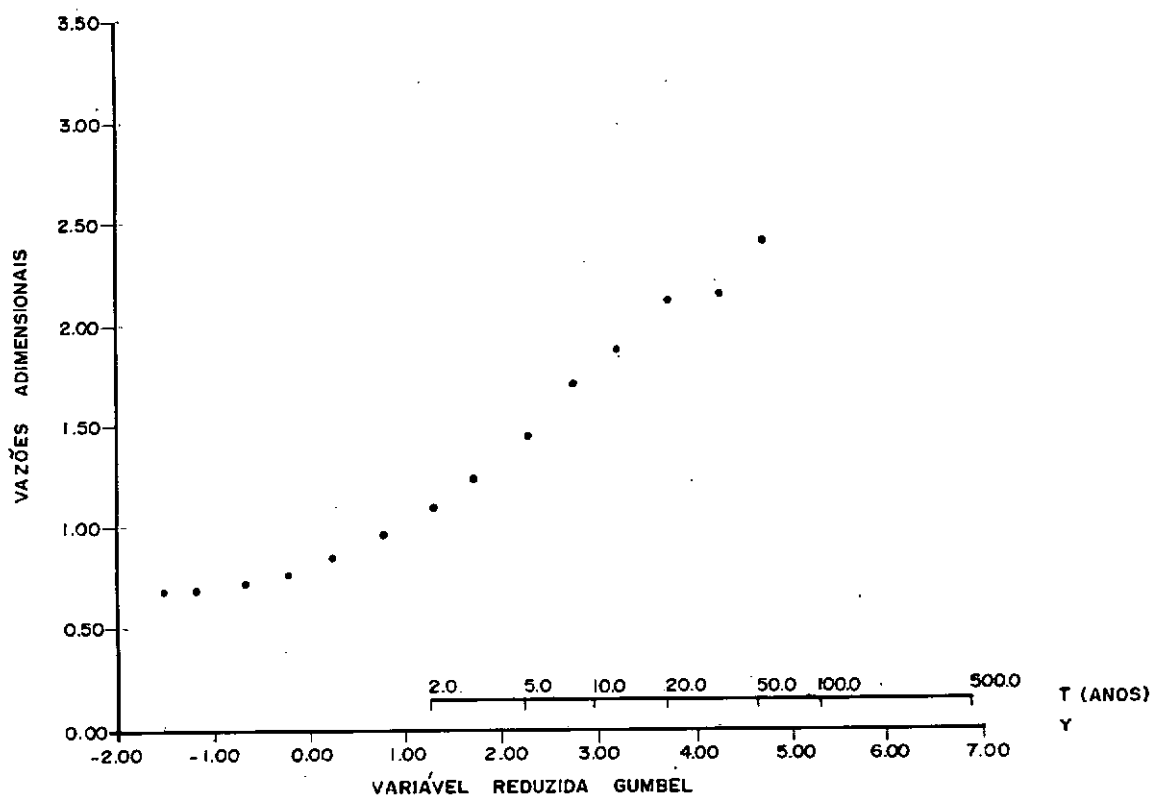


FIGURA 3.5 — DISTRIBUIÇÃO EMPÍRICA REGIONAL DAS SÉRIES PARCIAIS — SP 2 — BACIA DO RIO DOCE.

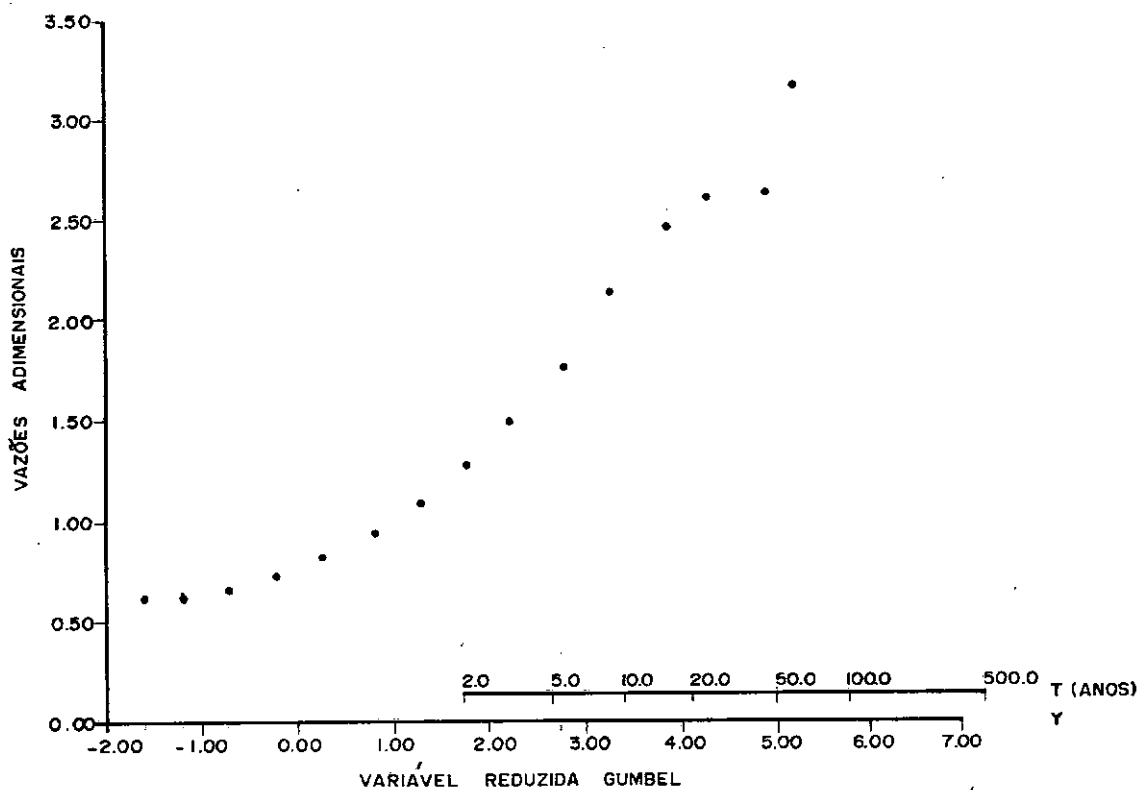


FIGURA 3.6 — DISTRIBUIÇÃO EMPÍRICA REGIONAL DAS SÉRIES PARCIAIS — SP 3 — BACIA DO RIO DOCE.

3.3.2-Homogeneidade da Bacia do Rio Doce

A utilização de informações regionais assume a hipótese de que a região considerada tem características homogêneas do ponto de vista hidrológico. Uma medida do grau de homogeneidade de uma região pode ser dada pela possibilidade de se ajustar uma distribuição de probabilidades única às séries de extremos de seus diferentes postos fluviométricos. À medida que a distribuição regional for considerada satisfatória para as diferentes séries de extremos da região, esta pode ser considerada homogênea.

Wallis (1980) propõe um método para a estimação dos parâmetros de uma distribuição Wakeby regional (ou central), que pode ser resumido em:

1. Ordenam-se as séries de extremos de cada posto no vetor (x_1, x_2, \dots, x_n) , com $x_1 < x_2 < x_3 < x_4 < \dots < x_n$
2. Calculam-se para cada posto as estimativas dos momentos ponderados pelas probabilidades, M_k , $k=0, 1, 2, 3, 4$:

$$M_k = \frac{1}{n} \sum_{i=1}^n x_i (1-F_i)^k \quad (3.27)$$

com

$$F_i = \frac{i - 0,35}{n} \quad (3.28)$$

3. Padronizam-se os M_K de cada posto fazendo $\bar{M}_K = M_K / M_0$. Se as séries de extremos já foram divididas pela sua média (séries adimensionais), esta etapa é dispensável.
4. Os valores regionais dos \bar{M}_K são definidos como a média dos \bar{M}_K de cada posto, ponderada pelo número de observações das respectivas séries.
5. Para a obtenção do conjunto de parâmetros (a, b, c, d, M) da distribuição Wakeby regional, utiliza-se o algoritmo designado por Landwehr et al.(1979) como "caminho $[V_m, m=0]$ ".
6. A distribuição Wakeby regional é dada então pela inversa de sua função de distribuição acumulada, ou seja:

$$q(T) = m + a \left[1 - \left(\frac{1}{\lambda T} \right)^b \right] - c \left[1 - \left(\frac{1}{\lambda T} \right)^{-d} \right] \quad (3.29)$$

sendo λ o número médio de picos por ano nas séries parciais, e fazendo $\lambda = 1$ nas séries de máximos anuais.

Foram definidas, com o auxílio deste método, as distribuições Wakeby regionais para quatro séries de extremos da bacia do rio Doce: SMA, SP1, SP2 e SP3. A tabela 3.6 mostra os valores dos parâmetros para as quatro séries de extremos, bem como os coeficientes de variação, assimetria e curtose das distribuições correspondentes. Estes coeficientes podem ser comparados com os valores amostrais apresentados na tabela 3.2.

SÉRIE	P A R Â M E T R O S					COEFICIENTES		
	a	b	c	d	m	CV	ASS	CUR
SMA	0,302649	10,7618	29,4716	0,0151178	0,270697	0,49	1,76	8,44
SP1	0,676193	205,896	2,76033	0,105939	0,00000	0,37	2,75	18,16
SP2	0,640747	1078,40	1,92791	0,157292	0,00000	0,44	3,61	33,48
SP3	0,623125	607,243	1,82435	0,171596	0,00000	0,47	3,88	41,46

Tabela 3.6 - Parâmetros das distribuições Wakeby regionais da bacia do rio Doce e seus coeficientes adimensionais.

A adequação do ajuste destas distribuições às séries históricas de cada posto foi verificada através do teste de Kolmogorov-Smirnov. A tabela 3.7 apresenta o número de postos cujas séries de extremos não foram bem ajustadas pelas distribuições regionais correspondentes. Verifica-se que o número de rejeições foi grande para as distribuições centrais das séries parciais, o que não ocorreu com as séries de máximos anuais, que foram bem ajustadas pela distribuição regional correspondente. Aparentemente, a bacia do rio Doce pode ser tratada como uma região hidrológicamente homogênea do ponto de vista das séries de máximos anuais o mesmo não acontecendo sob a ótica das séries parciais.

NÍVEL DE SIGNIFIC.	TAXA ANUAL DE PICOS			SÉRIE DE MÁXIMOS ANUAIS
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	
1%	6	9	12	1
5%	11	9	16	3
10%	12	13	18	3

Tabela 3.7 - Número de postos cujos ajustes à Wakeby central foram recusados pelo teste de Kolmogorov-Smirnov
Bacia do rio Doce - 29 postos

Esta aparente não homogeneidade das séries parciais da bacia do rio Doce poderia ser explicada pela utilização de uma distribuição regional inadequada. Com o objetivo de dirimir esta dúvida, foram ajustadas mais duas distribuições regionais às mesmas séries parciais: a distribuição generalizada de valores extremos (GEV, 3 parâmetros) e a distribuição exponencial (2 parâme

tros). Foram utilizadas as estatísticas regionais da tabela 3.2 para a estimação dos parâmetros destas duas distribuições. Novamente, os ajustes obtidos foram verificados pela aplicação do teste de Kolmogorov-Smirnov, e os resultados obtidos são mostrados nas tabelas 3.8 (distribuição de valores extremos) e 3.9 (distribuição exponencial). Estes resultados confirmaram o mau ajuste das distribuições regionais às séries parciais desta bacia.

α	TAXA ANUAL DE PICOS		
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$
1%	5	8	8
5%	7	12	18
10%	9	13	21

Tabela 3.8 - Número de postos cujos ajustes à distribuição de valores extremos regional foram recusados pelo teste de Kolmogorov-Smirnov.

Total de 29 postos - Bacia do rio Doce

α	TAXA ANUAL DE PICOS		
	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$
1%	6	8	12
5%	11	11	16
10%	11	12	20

Tabela 3.9 - Número de postos cujos ajustes à distribuição exponencial regional foram recusados no teste de Kolmogorov-Smirnov.

Total de 29 postos - Bacia do Rio Doce

Uma tentativa de levar em conta a influência da área de drenagem nestes resultados foi feita, ajustando-se distribuições centrais aos postos cujas áreas de drenagem estivessem compreendidas dentro de certos limites. Novamente, os resultados foram ruins, não tendo sido possível explicar o motivo do mau ajuste das distribuições regionais às séries parciais numa bacia onde a distribuição regional se adaptou muito bem às séries de máximos anuais.

CAPÍTULO IVSELEÇÃO DE MÉTODOS PARA ESTIMAÇÃO DE VAZÕES COM PERÍODOS DE
RECORRÊNCIA MUITO ELEVADOS4.1 - INTRODUÇÃO

Em projetos de obras hidráulicas de grande envergadura, em geral é necessário o cálculo de vazões com tempos de recorrência muito elevados. No caso de vertedores de grandes barragens, costuma-se trabalhar com vazões de tempo de recorrência da ordem de 1.000 a 10.000 anos. Existe um grande número de métodos para a estimação destes valores a partir de registros de vazões extremas. A aplicação destes métodos exige, inicialmente, a escolha da série a estudar (série de máximos anuais ou série parcial). Escolhida a série, é preciso determinar uma distribuição de probabilidades e o procedimento de estimação dos seus parâmetros. Ocorre que os resultados obtidos pela aplicação de diferentes métodos a um mesmo conjunto de dados podem diferir significativamente. A tabela 4.1 mostra, à guisa de exemplo, os resultados da aplicação de diversos modelos às séries de extremos do rio Doce em Colatina. É possível observar a grande variação dos resultados, principalmente na previsão de vazões com tempos de recorrência mais elevados.

Uma vez de posse destes resultados, deve-se determinar a vazão de projeto, podendo simplesmente optar pelo maior valor, pelo valor médio, ou, então, adotar algum procedimento para selecionar um dos métodos. No entanto, os procedimentos clássicos de seleção de modelos probabilísticos, como os testes de

	DISTRIBUIÇÃO	TEMPO DE RECORRÊNCIA (ANOS)							
		2	10	25	50	100	500	1.000	10.000
SÉRIE DE MÁXIMOS ANUAIS	GUMBEL	3.825	5.454	6.274	6.883	7.487	8.882	9.482	11.474
	LOG NORMAL	3.850	5.621	6.457	7.062	7.654	9.008	9.590	11.547
	EXPONENCIAL	3.667	5.451	6.467	7.236	8.005	9.789	10.558	13.111
	LOG PEARSON TIPO III	3.941	5.518	6.132	6.534	6.897	7.627	7.905	8.708
	WAKEBY	4.057	5.362	6.126	6.766	7.466	9.349	10.286	14.070
SÉRIE PARCIAL, $\lambda = 3$	GUMBEL	3.697	5.283	6.045	6.607	7.164	8.448	9.000	10.831
	LOG NORMAL	3.681	5.091	5.740	6.213	6.679	7.758	8.226	9.812
	EXPONENCIAL	3.593	5.620	6.640	7.397	8.148	9.884	10.630	13.108
	LOG PEARSON TIPO III	3.617	5.420	6.423	7.230	8.091	10.357	11.467	15.863
	WAKEBY	3.729	5.500	6.165	6.581	6.939	7.595	7.818	8.365

Tabela 4.1: Estimativas de Vazões (m^3/s) com Vários Tempos de Recorrência para o Rio Doce em Colatina usando Diferentes Modelos.

adequação de ajuste (ex: teste de Kolmogorov-Smirnov), neste caso não são conclusivos. Isto porque estes testes apenas avaliam a capacidade dos modelos de reproduzirem as frequências dos eventos registrados nos dados, nada podendo garantir, a princípio, sobre os seus desempenhos quando se trata da estimação de eventos menos frequentes. A aplicação destes testes se baseia na hipótese de que a distribuição de probabilidades que melhor reproduz as frequências amostrais será também a que melhor estimará as vazões com tempos de recorrência maiores. Naturalmente, esta hipótese só poderá ser verificada nos casos em que se tomarem amostras de uma população cuja distribuição de probabilidades seja conhecida.

Uma base objetiva para a seleção de um método de estimação de vazões com períodos de recorrência elevados pode ser estabelecida desde que se possa criar um cenário que represente o processo de ocorrência de vazões extremas a partir de um modelo matemático gerador de dados sintéticos. Dentro deste cenário, os valores corretos das vazões com períodos de recorrência milenares seriam conhecidos, e o desempenho de cada método poderia ser medido pelo correspondente erro cometido na estimação destes valores. Evidentemente, este cenário deve ser verossímil e incorporar características peculiares de cada área de interesse.

Uma possibilidade para a criação do cenário consiste simplesmente em produzir registros sintéticos de vazões extremas (séries de máximos anuais ou séries parciais), diretamente de uma distribuição de probabilidades. Para isto, naturalmente, é necessário encontrar-se uma distribuição capaz de produzir séries sintéticas com características semelhantes às séries registradas.

O uso da distribuição Wakeby para a criação de cenários deste tipo tem sido sugerido (Houghton, 1978; Landwear et alii, 1980; Kuczera, 1982), devido à sua grande versatilidade em reproduzir conjuntos de características amostrais (trata-se de uma distribuição com 5 parâmetros). Esta distribuição é definida através da inversa de sua função de distribuição, da seguinte forma:

$$Q = m + a \left[1 - (1-U)^b \right] - c \left[1 - (1-U)^{-d} \right] \quad (4.1)$$

onde Q é a variável aleatória Vazão Extrema de distribuição Wakeby U é a variável aleatória de distribuição uniforme no intervalo $[0,1]$, e $\phi = (m,a,b,c,d)$ é o conjunto de parâmetros que especifica a distribuição Wakeby. A partir da expressão 4.1 é possível sortear facilmente números aleatórios com distribuição Wakeby, o que permite a realização de experimentos simples de Monte - Carlo. Os valores de população para quaisquer tempos de recorrência T são obtidos diretamente desta expressão, fazendo-se:

$$q(T) = m + a \left[1 - \left(\frac{1}{T} \right)^b \right] - c \left[1 - \left(\frac{1}{T} \right)^{-d} \right] \quad (4.2)$$

Esta abordagem foi adotada por Damazio, Moreira, Costa e Kelman (1983) em um estudo aplicado às séries de máximos anuais da bacia do rio Doce. Neste estudo, foram ajustadas 8 distribuições Wakeby centrais (Wallis, 1980) às séries de vazões máximas anuais desta bacia para 8 diferentes durações. A principal conclusão deste trabalho foi a evidente superioridade das distribuições de dois parâmetros sobre distribuições mais complexas, no tocante à determinação de vazões com tempos de recorrência muito elevados. A distribuição exponencial (principalmente) e a distribuição Gumbel foram as mais eficientes neste sentido.

Muito embora esta abordagem seja conveniente para se avaliar a confiabilidade das extrapolações fornecidas pelo uso de uma dada distribuição de probabilidades, a comparação entre o uso de séries de máximos anuais ou séries parciais fica dificultada pela exigência adicional de se definirem simultaneamente conjuntos coerentes de parâmetros para as séries. Esta situação pode ser contornada através do uso de um modelo único, capaz de gerar hidrógrafas completas. De posse de um conjunto de hidrógrafas, aplicam-se os procedimentos usuais para a obtenção das séries parciais e de máximos anuais.

Uma alternativa para a obtenção de hidrógrafas sintéticas completas consiste no uso de modelos de simulação de bacias hidrográficas acoplados a modelos geradores de precipitação e evaporação potencial, sendo possível, desta forma, até mesmo a geração de vazões em diferentes pontos da bacia, simultaneamente. Estes modelos em geral exigem uma quantidade relativamente grande de informações que, quando em falta, dificultam a validação do modelo. Alternativamente, pode-se usar modelos estocásticos de vazões diárias (Barbosa, 1981), que utilizam apenas as informações dos registros de vazões diárias, dispensando os dados climático-meteorológicos. A desvantagem evidente destes dois tipos de modelo é o custo operacional de se gerar toda a hidrógrafa (a intervalo pelo menos diário) para a obtenção das séries de extremos. Este problema é tanto mais importante na medida que, em geral, não se conhece analiticamente a função que relaciona os parâmetros do modelo com as vazões milenares, sendo necessário o uso de simulação para este cálculo.

4.2 - Metodologia de Avaliação

O desempenho de cada método de estimação de vazões extremas deve ser avaliado a partir de índices que reflitam o seu comportamento médio, devendo estes índices caracterizar, principalmente, a dispersão dos resultados fornecidos por cada método, bem como sua tendência a subestimar ou superestimar os valores de população. Sendo este comportamento médio de difícil determinação teórica, propõe-se, a seguir, uma metodologia de avaliação baseada na geração de séries sintéticas de vazões extremas e no método de Monte Carlo.

1. Inicialmente, são geradas séries sintéticas de vazões diárias em número grande o suficiente para que se possa determinar empiricamente a distribuição de probabilidades dos máximos anuais, definindo-se assim os valores de população para as vazões com os diferentes tempos de recorrência desejados.
2. Estas séries são divididas em sub-amostras contendo N anos, sendo extraídas de cada uma as séries de vazões extremas a serem analisadas: séries de máximos anuais e séries parciais com diferentes taxas anuais de picos.
3. A cada uma destas séries de extremos são ajustadas as distribuições de probabilidades de interesse, sendo estimadas, a partir de cada ajuste obtido, as vazões para diferentes tempos de recorrência.

A repetição deste procedimento para todas as sub-amostras disponíveis fornece um conjunto de estimativas de $q(T)$ pa

ra cada distribuição ajustada a cada tipo de série de extremos. Os índices de desempenho de cada modelo são definidos a seguir, (Dalmazio et alii, 1983).

1. O erro médio absoluto da estimativa de $q(t)$ é calculado por:

$$EMA(T) = \frac{1}{NG} \sum_{i=1}^{NG} \left| \frac{\hat{q}_i(T) - q(T)}{q(T)} \right| \quad (4.3)$$

onde $\hat{q}_i(T)$ é a estimativa feita a partir da i -ésima sub-amostra, $q(T)$ é o valor de população, e NG é o número total de sub-amostras disponíveis. A precisão na estimativa deste índice $EMA(T)$ pode ser medida a partir do seu desvio padrão, dado por:

$$SE(T) = \sqrt{\frac{1}{NG^2} \sum_{i=1}^{NG} \left[\left| \frac{\hat{q}_i(T) - q(T)}{q(T)} \right| - EMA(T) \right]^2} \quad (4.4)$$

2. O valor esperado de $\hat{q}_i(T)/q(T)$ dá uma idéia da tendência do estimador de $q(T)$ e pode ser calculado por:

$$MED(T) = \frac{1}{NG} \sum_{i=1}^{NG} \frac{\hat{q}_i(T)}{q(T)} \quad (4.5)$$

Da mesma forma, a precisão da estimativa de $MED(T)$ é dada pelo seu desvio padrão:

$$SM(T) = \sqrt{\frac{1}{NG^2} \sum_{i=1}^{NG} \left[\frac{\hat{q}_i(T)}{q(T)} - MED(T) \right]^2} \quad (4.6)$$

3. As probabilidades de subestimação ou superestimação dos valores de $q(T)$ são:

$$p_1 = P \left[\hat{x}_i(T) \leq 0,9 \right] \quad (4.7)$$

$$p_2 = P \left[0,9 \leq \hat{x}_i(T) \leq 1,0 \right] \quad (4.8)$$

$$p_3 = P \left[1,0 \leq \hat{x}_i(T) \leq 1,1 \right] \quad (4.9)$$

$$p_4 = P \left[\hat{x}_i(T) \geq 1,1 \right] \quad (4.10)$$

$$\text{com } \hat{x}_i(T) = \frac{\hat{q}_i(T)}{q(T)} \quad (4.11)$$

Estas probabilidades podem ser estimadas empiricamente, por simples enumeração.

Para se avaliar a influência do tamanho dos registros na precisão das estimativas de vazões extremas, estes índices devem ser levantados para diferentes tamanhos de sub-amostras.

O método que fornece os menores valores de EMA(T) é considerado o melhor, sendo os valores de MED(T) e das probabilidades de sub e superestimação resultados auxiliares na avaliação.

4.3 - Aplicação a Séries Sintéticas - Modelo Xadrez

A metodologia descrita no item anterior foi aplicada a uma série sintética de vazões diárias contendo 10.000 anos hidrológicos. Estas séries foram geradas através de um modelo estocástico não paramétrico de vazões diárias desenvolvido por Jerson Kelman (comunicação pessoal), o modelo XADREZ, cuja descrição é

suscintamente apresentada no apêndice B.

O modelo foi calibrado para o registro de vazões médias diárias do rio Doce em Colatina. Este posto fluviométrico está situado próximo à usina hidrelétrica de Mascarenhas, tendo sua bacia de contribuição uma área de 76.615 km². O registro disponível para este estudo começa em 1938 e termina em 1973, proporcionando um total de 29 anos hidrológicos sem falhas. Foram considerados na modelagem apenas os dias correspondentes ao período chuvoso da bacia do rio Doce (1º de outubro a 30 de abril), num total de 212 dias.

Foram gerados 100.000 anos de vazões diárias, sendo que apenas 10.000 destas séries foram integralmente armazenadas para a aplicação da metodologia do item 4.2. As 90.000 séries restantes foram utilizadas apenas para a determinação da distribuição empírica dos máximos anuais.

As características das séries geradas são comparadas com as das séries históricas no item 4.3.1, onde também é apresentada a distribuição empírica dos máximos anuais gerados, bem como os valores de população para os diferentes tempos de recorrência a serem tratados na análise subsequente. Os resultados da aplicação da metodologia são apresentados e discutidos em 4.3.2.

Com o objetivo de se investigar até que ponto os resultados obtidos são específicos para o modelo de geração usado, foram geradas novas séries sintéticas, através de outro modelo de geração, o modelo DIANA, (Kelman et al. 1983). Este modelo,

ao contrário do XADREZ, traz em sua formulação uma concepção do processo de ocorrência de vazões, tratando de forma diferenciada as ascensões e recessões da hidrografa. Os resultados obtidos são apresentados e discutidos no item 4.4.

4.3.1-Calibração do Modelo Xadrez

A avaliação da qualidade das séries geradas pelo modelo XADREZ é feita através da comparação de algumas propriedades das séries históricas com as correspondentes propriedades das séries geradas, de acordo com metodologia descrita em Kelman, Damazio e Costa (1983). A abordagem adotada para esta comparação consiste em calcular, para a série histórica e para diversas séries geradas com o mesmo comprimento do registro histórico, algumas estatísticas específicas, ou índices. Resultam daí distribuições empíricas de probabilidades destes índices para as séries geradas. Uma análise da posição dos índices obtidos para as séries históricas dentro destas distribuições empíricas permite avaliar se a série histórica poderia ter sido sorteada da mesma população que as séries sintéticas.

Os índices examinados foram: a média, o desvio padrão, a assimetria e a curtose das vazões médias diárias; estes mesmos índices para as vazões máximas anuais; a máxima vazão observada; e o volume de espera necessário para regularizar uma vazão igual à média dos máximos anuais.

A tabela 4.2 apresenta os índices da série histórica e os índices médios das séries geradas pelo modelo XADREZ. Pa-

Í N D I C E		VAZÕES MÉDIAS DIÁRIAS				VAZÕES MÁXIMAS ANUAIS				VALOR MÁXIMO OBSERVADO	VOLUME DE ESPERA
		Q	S	ASS	CUR	Q	S	ASS	CUR		
SÉRIE HISTÓRICA		1204	863	1,87	7,48	4006	1090	0,11	2,56	6405	$0,17 \times 10^5$
SÉRIES SINTÉTICAS	POPULAÇÃO	1237	977	2,25	10,86	3969	1408	1,02	5,12	11697	$0,30 \times 10^6$
	MÍNIMO	1096	811	1,52	5,75	3442	993	-0,32	1,73	5669	$0,10 \times 10^5$
	MÉDIO	1237	970	2,13	9,70	3972	1367	0,78	4,06	7745	$0,70 \times 10^5$
	MÁXIMO	1416	1222	3,20	23,40	4541	1888	3,10	14,65	11697	$0,30 \times 10^6$
P{SINT > HIST}		0,76	0,88	0,74	0,74	0,32	0,88	0,85	0,74	0,88	0,91

Tabela 4,2 - Índices das séries sintéticas gerados pelo modelo XADRES e da série histórica do rio Doce em Colatina. 34 séries sintéticas de 29 anos.

ra permitir uma avaliação da semelhança entre as duas séries, aparecem ainda nesta tabela as faixas de variação dos índices das séries geradas, e as probabilidades de que estes índices sejam superiores aos da série histórica. A hipótese de que os máximos anuais históricos foram sorteados das séries geradas não pode ser recusada a um nível de significância de 10% pelo teste do chi-quadrado.

A Figura 4.1 apresenta as vazões médias de cada dia nas duas séries. Observa-se que as séries geradas mantêm a sazonalidade das séries históricas. A figura 4.2 mostra as hidrógrafas com máximos anuais medianos das séries geradas e históricas. Aparentemente, a cheia mediana das séries geradas não tem aspecto inverossímil, se comparada com a correspondente do histórico.

A Figura 4.3 mostra a distribuição empírica dos máximos anuais, estando marcados na mesma os pontos correspondentes aos máximos anuais do registro histórico. Desta figura são obtidos os valores de população para diferentes tempos de recorrência que servirão de base para a aplicação da metodologia de avaliação dos métodos de estimação de vazões extremas. A tabela 4.3 apresenta estes valores de população.

T	2	10	25	50	100	500	1000	10000
q(T)	3745	5869	6994	7812	8639	10436	11237	13228

Tabela 4.3 - Valores da População de Máximos Anuais gerada pelo Modelo XADREZ para Colatina.

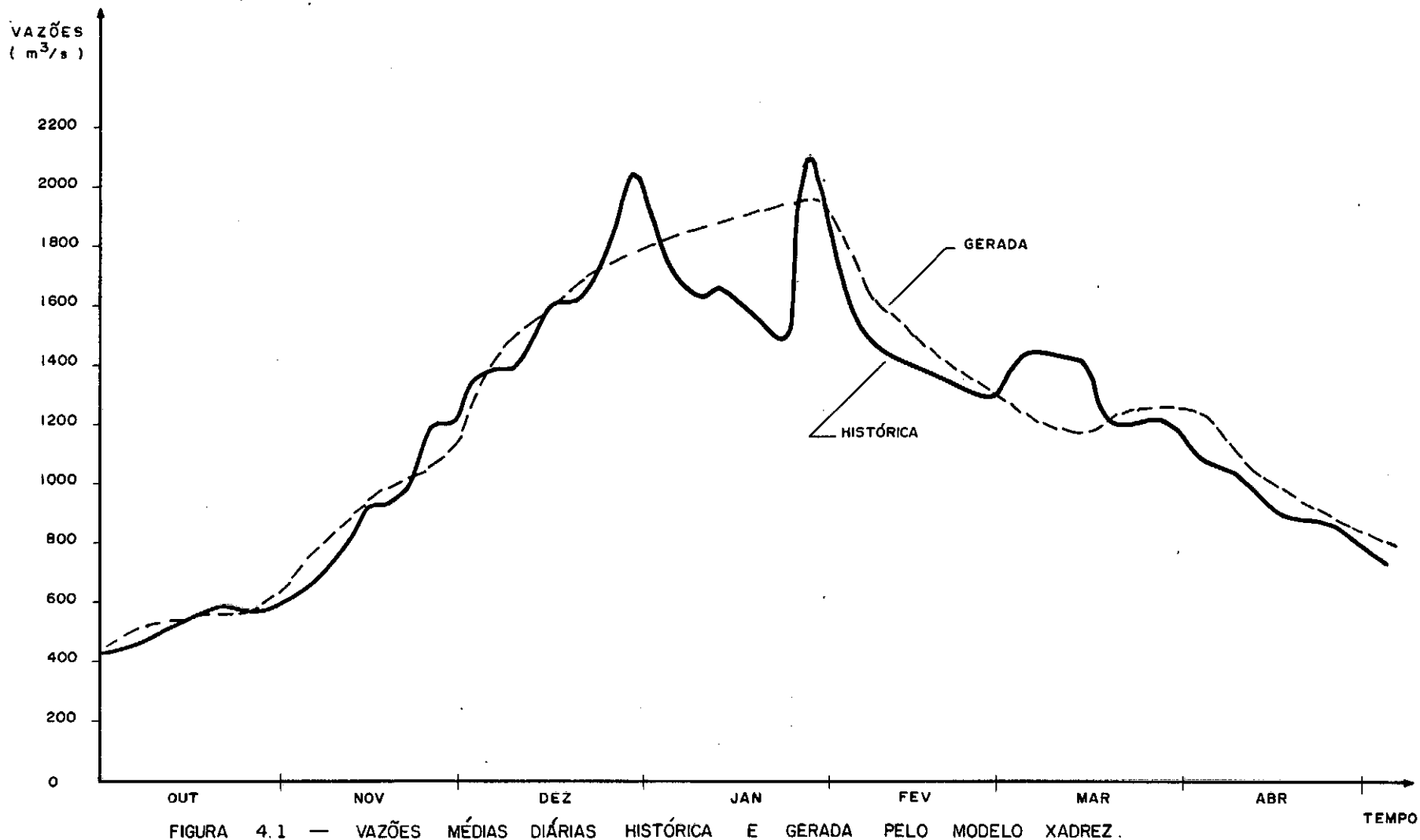
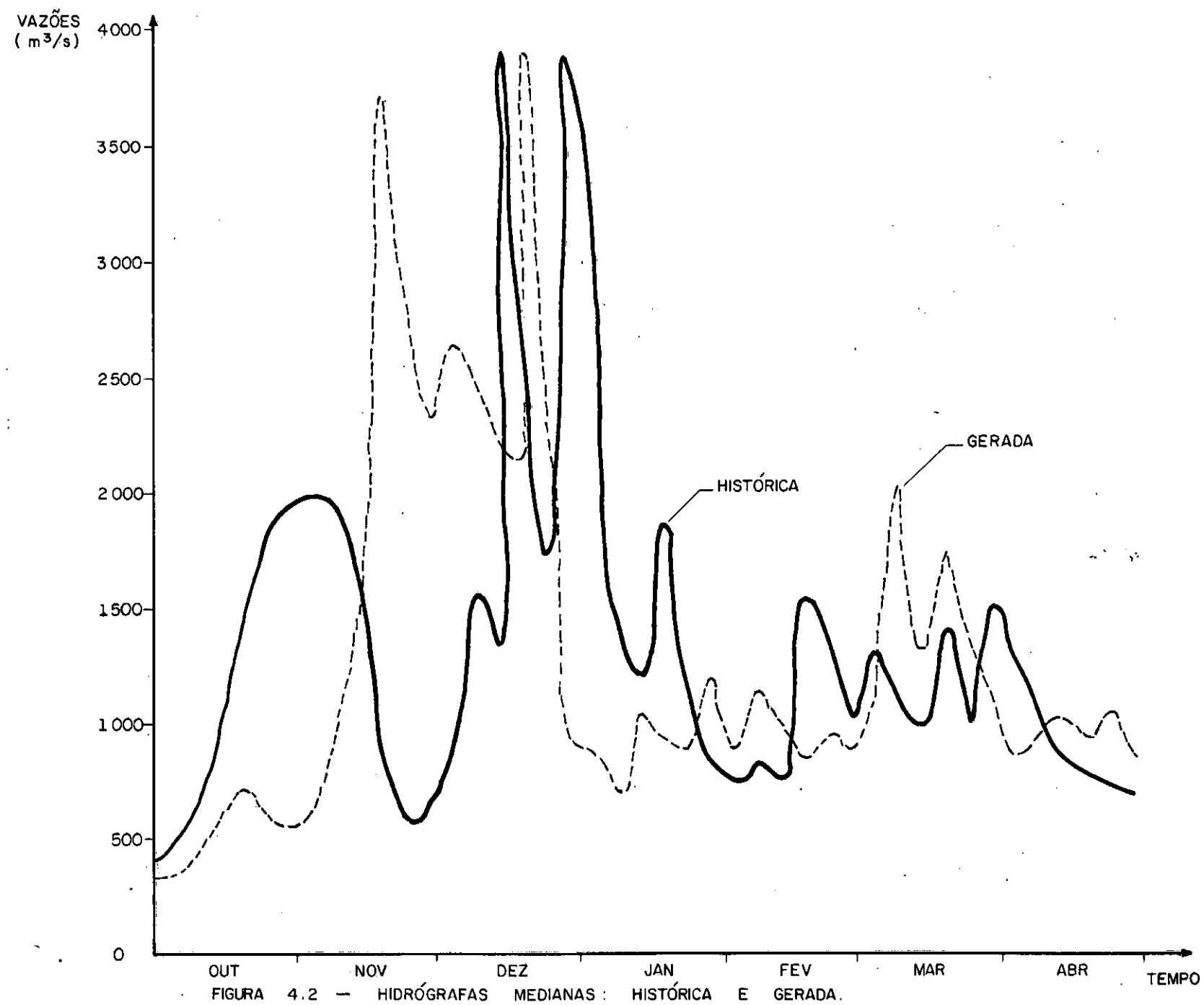


FIGURA 4.1 — VAZÕES MÉDIAS DIÁRIAS HISTÓRICA E GERADA PELO MODELO XADREZ.



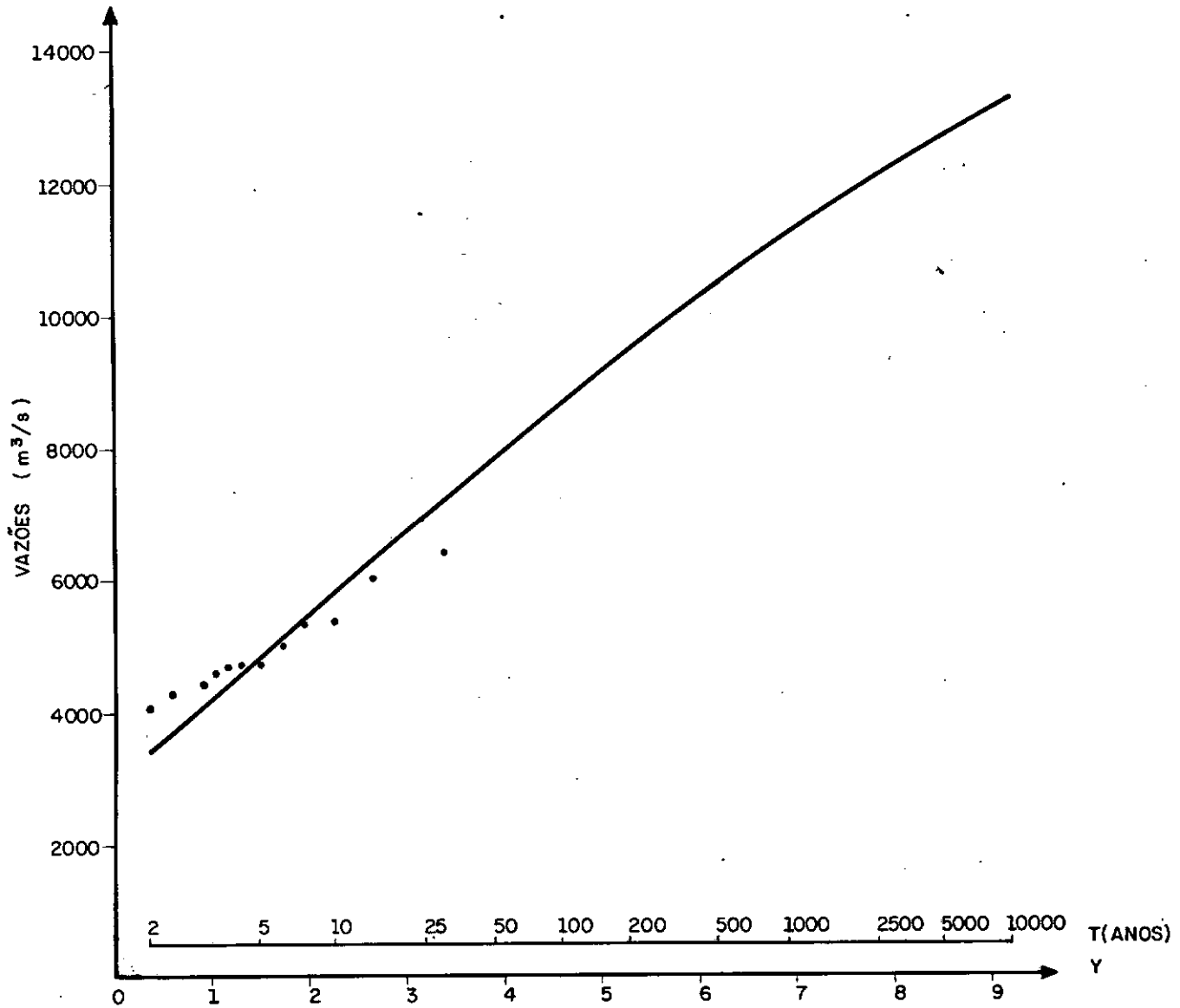


FIGURA 4.3 — DISTRIBUIÇÃO EMPÍRICA DOS MÁXIMOS ANUAIS GERADOS PELO MODELO XADREZ — RIO DOCE EM COLATINA.

4.3.2-Máximos Anuais x Séries Parciais

A metodologia descrita no item 4.2 foi aplicada às 10000 séries sintéticas de vazões diárias disponíveis, tendo sido estudadas quatro diferentes séries de vazões extremas: SMA, SP1, SP2 e SP3. Foram escolhidas para aplicação as dez distribuições de probabilidades teóricas seguintes:

1. distribuição normal (N).
2. distribuição log-normal de 2 parâmetros (LN2).
3. distribuição log-normal de 3 parâmetros (LN3).
4. distribuição exponencial de 2 parâmetros (EXP).
5. distribuição gama de 2 parâmetros (GAM2).
6. distribuição gama de 3 parâmetros ou Pearson tipo 3 (P3).
7. distribuição log-Pearson tipo 3 (LP3).
8. distribuição Gumbel (GU).
9. distribuição generalizada de valores extremos (GEV).
10. distribuição Wakeby (WAK).

Estas distribuições foram ajustadas pelo método dos momentos, à exceção da distribuição Wakeby, ajustada pelo método dos momentos ponderados pelas probabilidades (Landwehr et al., 1979), e da distribuição exponencial, ajustada tanto pelo método dos momentos, quanto pelo método da máxima verossimilhança. No apêndice A são apresentados, de forma sumária, os métodos de estimação de parâmetros e as expressões usadas para a estimação dos quantis para cada distribuição. Para a consideração da influência do ta-

manho das amostras no desempenho de cada método, foram tomadas, das 1000 séries de extremos sub-amostras com $N=5, 10, 15, 20, 25, 30, 40, 50$ e 60 anos.

As séries parciais foram obtidas pelo critério "a" descrito em 2.3.2, ou seja, foi fixado um nível de truncamento e considerado como pico a máxima vazão de cada período de excesso. É computacionalmente muito trabalhoso procurar, para cada sub-amostra, o nível de truncamento que fornece exatamente $N, 2N$ ou $3N$ picos para a obtenção de $SP1, SP2$ e $SP3$, respectivamente. Assim sendo, foi fixado como nível de truncamento a metade da média dos máximos anuais de cada sub-amostra e, a partir deste nível, definida uma série parcial básica. As séries $SP1, SP2$ e $SP3$ foram então obtidas, selecionando-se os $N, 2N$ e $3N$ maiores valores das séries parciais básicas de cada sub-amostra. Este procedimento permitiu simplificar bastante a parte computacional.

As tabelas 4.4 a 4.7 e 4.8 a 4.11 apresentam uma saída de resultados típica, obtida para as sub-amostras de 20 anos. Foram analisadas as previsões de vazões com tempos de recorrência pequenos (2, 10, 25 e 50 anos) e grandes (100, 500, 1000 e 10000 anos). Estas tabelas mostram, para cada tempo de recorrência T e para cada distribuição de probabilidades, os valores (da esquerda para a direita): $MED(T)$ e $SM(T)$, $EMA(T)$ e $SE(T)$, p_1, p_2, p_3 e p_4 (equações 4.3 a 4.10). Naturalmente, a apresentação de todos os resultados assim detalhados seria problemática. Assim sendo, os resultados obtidos são sumarizados em tabelas-resumo, que são descritas a seguir.

VAZÕES DE PROJETO PARA O POSTO DE COLATINA OBTIDAS PELO METODO DE MAXIMOS ANUAIS

N= 20 NG= 500

ME	TEMPO DE RETORNO= 100.								TEMPO DE RETORNO= 500.							
	POPULACAO= 8639.0								POPULACAO= 10436.0							
TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1			MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1		
N	0.94	0.00	0.17	0.00	0.728	0.192	0.070	0.010	0.77	0.00	0.23	0.00	0.878	0.102	0.018	0.002
LN	0.99	0.01	0.13	0.00	0.300	0.242	0.222	0.236	1.01	0.01	0.15	0.01	0.320	0.204	0.186	0.290
LN3	0.86	0.01	0.14	0.01	0.594	0.196	0.128	0.082	0.82	0.01	0.22	0.01	0.678	0.172	0.078	0.072
EXP	1.05	0.01	0.13	0.00	0.172	0.222	0.248	0.358	1.09	0.01	0.15	0.01	0.142	0.190	0.222	0.446
*EXP	1.37	0.01	0.38	0.01	0.006	0.020	0.078	0.896	1.47	0.01	0.48	0.01	0.006	0.008	0.042	0.444
GAM	0.93	0.01	0.13	0.00	0.450	0.268	0.168	0.114	0.90	0.01	0.15	0.00	0.552	0.216	0.130	0.100
P3	0.88	0.01	0.21	0.01	0.468	0.200	0.160	0.172	0.86	0.01	0.25	0.01	0.532	0.172	0.106	0.190
LP3	0.60	0.02	0.51	0.02	0.612	0.078	0.110	0.200	0.61	0.02	0.57	0.02	0.610	0.048	0.072	0.270
GU	0.97	0.01	0.11	0.00	0.310	0.276	0.232	0.182	0.98	0.01	0.12	0.00	0.326	0.260	0.222	0.192
GEV	0.93	0.01	0.15	0.00	0.472	0.202	0.172	0.154	0.91	0.01	0.20	0.01	0.532	0.172	0.098	0.198
WAK	1.04	0.01	0.16	0.01	0.258	0.206	0.190	0.346	1.15	0.02	0.29	0.01	0.276	0.138	0.124	0.462

ME	TEMPO DE RETORNO= 1000.								TEMPO DE RETORNO= 10000.							
	POPULACAO= 11237.0								POPULACAO= 13228.0							
TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1			MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1		
N	0.74	0.00	0.26	0.00	0.922	0.070	0.006	0.002	0.70	0.00	0.30	0.00	0.970	0.028	0.002	0.000
LN	1.01	0.01	0.16	0.01	0.330	0.192	0.184	0.294	1.08	0.01	0.21	0.01	0.274	0.148	0.174	0.404
LN3	0.80	0.01	0.24	0.01	0.708	0.164	0.058	0.070	0.79	0.01	0.27	0.01	0.728	0.114	0.068	0.090
EXP	1.10	0.01	0.16	0.01	0.138	0.186	0.212	0.464	1.18	0.01	0.21	0.01	0.068	0.134	0.172	0.626
*EXP	1.50	0.01	0.51	0.01	0.006	0.008	0.034	0.352	1.66	0.01	0.66	0.01	0.002	0.004	0.010	0.984
GAM	0.88	0.01	0.16	0.00	0.596	0.190	0.122	0.092	0.88	0.01	0.18	0.01	0.600	0.168	0.122	0.110
P3	0.85	0.02	0.27	0.01	0.548	0.160	0.092	0.200	0.86	0.02	0.30	0.01	0.536	0.144	0.082	0.238
LP3	0.62	0.03	0.60	0.02	0.608	0.040	0.068	0.284	0.71	0.03	0.72	0.02	0.572	0.040	0.030	0.358
GU	0.98	0.01	0.13	0.00	0.334	0.258	0.208	0.200	1.02	0.01	0.14	0.00	0.256	0.222	0.216	0.304
GEV	0.90	0.01	0.22	0.01	0.560	0.152	0.090	0.198	0.92	0.02	0.29	0.01	0.568	0.122	0.074	0.236
WAK	1.22	0.02	0.37	0.02	0.282	0.116	0.102	0.500	1.69	0.06	0.86	0.05	0.262	0.082	0.054	0.582

ORS: EM 100% DOS GRUPOS O METODO WAK CONVERGIU

Tabela 4.4

VAZÕES DE PROJETO PARA O POSTO DE COLATINA OBTIDAS COM 1 PICOS POR ANO

N= 20 NG= 500

TEMPO DE RETORNO= 100. POPULACAO= 8639.0																TEMPO DE RETORNO= 500. POPULACAO= 10436.0																																																																																																																																																																										
ME	MED				EMA				<0.9				0.9-1.0				1.0-1.1				>1.1				ME	MED				EMA				<0.9				0.9-1.0				1.0-1.1				>1.1																																																																																																																																												
TUDO																									TUDO																																																																																																																																																																	
N	0.80	0.00	0.20	0.00	0.804	0.158	0.032	0.006	0.72	0.00	0.28	0.00	0.934	0.060	0.006	0.000	LN	0.83	0.01	0.18	0.00	0.702	0.204	0.074	0.020	0.79	0.01	0.22	0.01	0.796	0.136	0.058	0.010	LN3	0.88	0.01	0.17	0.00	0.614	0.194	0.110	0.082	0.84	0.01	0.21	0.01	0.684	0.144	0.088	0.084	EXP	0.97	0.01	0.13	0.00	0.342	0.256	0.210	0.192	0.97	0.01	0.15	0.00	0.352	0.230	0.194	0.224	*EXP	0.98	0.01	0.11	0.00	0.274	0.298	0.222	0.206	0.99	0.01	0.13	0.00	0.292	0.246	0.214	0.248	GAM	0.85	0.01	0.17	0.00	0.658	0.200	0.096	0.046	0.80	0.01	0.22	0.01	0.782	0.126	0.068	0.024	P3	0.93	0.01	0.15	0.00	0.478	0.206	0.162	0.154	0.92	0.01	0.19	0.01	0.520	0.170	0.106	0.204	LP3	0.96	0.01	0.16	0.01	0.394	0.216	0.166	0.224	1.02	0.01	0.23	0.01	0.382	0.144	0.142	0.332	GU	0.91	0.01	0.14	0.00	0.526	0.234	0.144	0.096	0.89	0.01	0.16	0.00	0.574	0.208	0.126	0.092	GEV	0.92	0.01	0.15	0.00	0.494	0.218	0.154	0.134	0.92	0.01	0.19	0.01	0.512	0.178	0.104	0.206	WAK	1.06	0.01	0.17	0.01	0.236	0.194	0.208	0.362	1.20	0.02	0.31	0.01	0.220	0.152	0.120	0.508

TEMPO DE RETORNO= 1000. POPULACAO= 11237.0																TEMPO DE RETORNO= 10000. POPULACAO= 13228.0																																																																																																																																																																										
ME	MED				EMA				<0.9				0.9-1.0				1.0-1.1				>1.1				ME	MED				EMA				<0.9				0.9-1.0				1.0-1.1				>1.1																																																																																																																																												
TUDO																									TUDO																																																																																																																																																																	
N	0.69	0.00	0.31	0.00	0.968	0.028	0.004	0.000	0.64	0.00	0.36	0.00	0.990	0.008	0.002	0.000	LN	0.77	0.01	0.24	0.01	0.840	0.104	0.046	0.010	0.76	0.01	0.25	0.01	0.818	0.106	0.056	0.020	LN3	0.83	0.01	0.23	0.01	0.718	0.116	0.080	0.086	0.83	0.01	0.26	0.01	0.694	0.104	0.090	0.122	EXP	0.97	0.01	0.15	0.00	0.362	0.222	0.184	0.232	1.02	0.01	0.17	0.01	0.292	0.214	0.150	0.344	*EXP	0.99	0.01	0.13	0.00	0.296	0.240	0.208	0.256	1.05	0.01	0.15	0.01	0.216	0.206	0.218	0.360	GAM	0.77	0.01	0.24	0.01	0.808	0.116	0.054	0.022	0.75	0.01	0.27	0.01	0.824	0.100	0.052	0.024	P3	0.92	0.01	0.21	0.01	0.536	0.152	0.106	0.206	0.96	0.01	0.24	0.01	0.498	0.146	0.086	0.270	LP3	1.06	0.02	0.27	0.01	0.374	0.130	0.118	0.378	1.31	0.03	0.48	0.03	0.272	0.098	0.092	0.538	GU	0.88	0.01	0.17	0.00	0.582	0.214	0.114	0.090	0.90	0.01	0.17	0.01	0.554	0.182	0.136	0.128	GEV	0.93	0.01	0.21	0.01	0.518	0.158	0.096	0.228	1.01	0.02	0.28	0.01	0.474	0.108	0.090	0.328	WAK	1.29	0.02	0.41	0.02	0.220	0.114	0.122	0.544	1.85	0.06	0.97	0.06	0.210	0.058	0.058	0.674

OBS: EM 100% DOS GRUPOS O METODO WAK CONVERGIU

Tabela. 4.5

VAZÕES DE PROJETO PARA O POSTO DE COLATINA OBTIDAS COM 2 PICOS POR ANO

N= 20 NG= 500

ME	TEMPO DE RETORNO= 100.								TEMPO DE RETORNO= 500.							
	TODO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1		TODO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	
N	0.75	0.00	0.25	0.00	0.958	0.040	0.002	0.000	0.67	0.00	0.33	0.00	0.994	0.006	0.000	0.000
LN	0.80	0.00	0.20	0.00	0.834	0.134	0.032	0.000	0.76	0.00	0.24	0.00	0.888	0.090	0.020	0.002
LN3	0.87	0.01	0.17	0.00	0.644	0.194	0.096	0.066	0.84	0.01	0.20	0.01	0.690	0.146	0.086	0.078
EXP	0.97	0.01	0.11	0.00	0.312	0.290	0.234	0.164	0.98	0.01	0.12	0.00	0.316	0.262	0.230	0.192
*EXP	0.97	0.00	0.09	0.00	0.300	0.334	0.232	0.134	0.97	0.01	0.10	0.00	0.298	0.300	0.244	0.158
GAM	0.83	0.01	0.19	0.00	0.742	0.174	0.070	0.014	0.78	0.01	0.23	0.01	0.842	0.110	0.036	0.012
P3	0.95	0.01	0.14	0.00	0.424	0.234	0.174	0.168	0.95	0.01	0.17	0.01	0.454	0.190	0.126	0.230
LP3	1.03	0.01	0.15	0.01	0.272	0.228	0.164	0.336	1.15	0.01	0.25	0.01	0.216	0.128	0.156	0.500
GU	0.89	0.01	0.13	0.00	0.554	0.276	0.128	0.042	0.87	0.01	0.15	0.00	0.602	0.238	0.120	0.040
GEV	0.94	0.01	0.14	0.00	0.444	0.234	0.174	0.148	0.96	0.01	0.17	0.01	0.424	0.196	0.122	0.258
WAK	1.06	0.01	0.17	0.01	0.248	0.196	0.186	0.370	1.19	0.02	0.32	0.01	0.252	0.174	0.118	0.506

ME	TEMPO DE RETORNO= 1000.								TEMPO DE RETORNO= 10000.							
	TODO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1		TODO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	
N	0.65	0.00	0.35	0.00	0.998	0.002	0.000	0.000	0.60	0.00	0.40	0.00	1.000	0.000	0.000	0.000
LN	0.75	0.01	0.25	0.00	0.894	0.086	0.018	0.002	0.75	0.01	0.25	0.01	0.874	0.082	0.038	0.006
LN3	0.83	0.01	0.22	0.01	0.702	0.126	0.088	0.084	0.86	0.01	0.24	0.01	0.666	0.112	0.080	0.142
EXP	0.98	0.01	0.12	0.00	0.318	0.260	0.226	0.196	1.03	0.01	0.14	0.00	0.232	0.228	0.210	0.330
*EXP	0.97	0.01	0.10	0.00	0.306	0.294	0.240	0.160	1.02	0.01	0.11	0.00	0.192	0.260	0.266	0.282
GAM	0.76	0.01	0.25	0.01	0.866	0.100	0.028	0.006	0.74	0.01	0.26	0.01	0.874	0.088	0.026	0.012
P3	0.96	0.01	0.18	0.01	0.460	0.180	0.126	0.234	1.01	0.01	0.22	0.01	0.400	0.160	0.132	0.308
LP3	1.22	0.02	0.31	0.01	0.202	0.098	0.154	0.546	1.60	0.03	0.65	0.03	0.116	0.060	0.062	0.762
GU	0.86	0.01	0.16	0.00	0.628	0.222	0.110	0.040	0.89	0.01	0.15	0.00	0.556	0.246	0.126	0.072
GEV	0.98	0.01	0.19	0.01	0.418	0.174	0.122	0.286	1.11	0.02	0.29	0.01	0.336	0.118	0.118	0.428
WAK	1.26	0.02	0.41	0.02	0.266	0.108	0.088	0.538	1.79	0.06	0.95	0.06	0.262	0.064	0.056	0.618

DRS: EM 100% DOS GRUPOS O METODO WAK CONVERGIU

Tabela 4.6

VAZÕES DE PROJETO PARA O POSTO DE COLATINA OBTIDAS COM 3 PICOS POR ANO

N= 20 NG= 440

TEMPO DE RETORNO= 100. POPULACAO= 8639.0								TEMPO DE RETORNO= 500. POPULACAO= 10436.0							
ME	TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	ME	TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1
N	0.71	0.00	0.29	0.00	0.991	0.009	0.000	0.64	0.00	0.36	0.00	1.000	0.000	0.000	0.000
LN	0.78	0.00	0.22	0.00	0.914	0.086	0.007	0.75	0.00	0.25	0.00	0.930	0.064	0.007	0.000
LN3	0.86	0.01	0.17	0.00	0.659	0.193	0.091	0.84	0.01	0.20	0.01	0.682	0.150	0.095	0.071
FXP	0.96	0.01	0.10	0.00	0.307	0.336	0.225	0.97	0.01	0.11	0.00	0.318	0.307	0.218	0.157
*EXP	0.93	0.00	0.10	0.00	0.414	0.364	0.182	0.93	0.00	0.10	0.00	0.420	0.350	0.189	0.041
GAM	0.81	0.00	0.20	0.00	0.814	0.150	0.032	0.76	0.01	0.24	0.00	0.889	0.091	0.020	0.000
P3	0.96	0.01	0.13	0.00	0.400	0.241	0.184	0.97	0.01	0.16	0.01	0.416	0.211	0.139	0.234
LP3	1.08	0.01	0.16	0.01	0.180	0.193	0.205	1.24	0.01	0.30	0.01	0.105	0.123	0.136	0.636
GU	0.87	0.00	0.14	0.00	0.620	0.257	0.109	0.86	0.01	0.15	0.00	0.680	0.214	0.093	0.014
GEV	0.95	0.01	0.13	0.00	0.402	0.252	0.193	0.99	0.01	0.16	0.01	0.348	0.218	0.150	0.284
WAK	1.04	0.01	0.17	0.01	0.259	0.214	0.157	1.14	0.02	0.30	0.01	0.298	0.118	0.120	0.464

TEMPO DE RETORNO= 1000. POPULACAO= 11237.0								TEMPO DE RETORNO= 10000. POPULACAO= 13228.0							
ME	TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	ME	TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1
N	0.61	0.00	0.39	0.00	1.000	0.000	0.000	0.57	0.00	0.43	0.00	1.000	0.000	0.000	0.000
LN	0.73	0.00	0.27	0.00	0.934	0.061	0.005	0.74	0.01	0.26	0.01	0.909	0.066	0.025	0.000
LN3	0.84	0.01	0.21	0.01	0.700	0.132	0.093	0.87	0.01	0.24	0.01	0.645	0.114	0.086	0.155
FXP	0.97	0.01	0.11	0.00	0.325	0.298	0.218	1.01	0.01	0.11	0.00	0.220	0.261	0.234	0.284
*EXP	0.92	0.00	0.10	0.00	0.434	0.343	0.182	0.96	0.01	0.09	0.00	0.275	0.361	0.252	0.111
GAM	0.75	0.01	0.26	0.01	0.902	0.082	0.016	0.73	0.01	0.27	0.01	0.900	0.080	0.020	0.000
P3	0.97	0.01	0.17	0.01	0.427	0.193	0.125	1.04	0.01	0.20	0.01	0.330	0.191	0.143	0.336
LP3	1.33	0.02	0.38	0.02	0.084	0.102	0.102	1.83	0.03	0.85	0.03	0.045	0.025	0.057	0.873
GU	0.85	0.01	0.16	0.00	0.693	0.209	0.084	0.87	0.01	0.15	0.00	0.623	0.236	0.114	0.027
GEV	1.01	0.01	0.18	0.01	0.339	0.189	0.155	1.17	0.02	0.30	0.01	0.241	0.132	0.123	0.505
WAK	1.20	0.02	0.37	0.02	0.309	0.100	0.100	1.59	0.06	0.78	0.06	0.305	0.068	0.057	0.570

OBS: EM 100% DOS GRUPOS O METODO WAK CONVERGIU

Tabela 4.7

VAZÕES DE PROJETO PARA O POSTO DE COLATINA OBTIDAS PELO MÉTODO DE MÁXIMOS ANUAIS

N= 20 NG= 500

1. TEMPO DE RETORNO= 2. POPULAÇÃO= 3745.0										10. TEMPO DE RETORNO= 10. POPULAÇÃO= 5869.0									
ME	3. TEMPO DE RETORNO=				4. TEMPO DE RETORNO=				ME	11. TEMPO DE RETORNO=				12. TEMPO DE RETORNO=					
TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	
N	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
LN	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
LN3	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
EXP	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
*EXP	0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
GAM	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
P3	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
LP3	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
GH	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
GEV	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
*AK	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

25. TEMPO DE RETORNO= 25. POPULAÇÃO= 6994.0										50. TEMPO DE RETORNO= 50. POPULAÇÃO= 7812.0									
ME	3. TEMPO DE RETORNO=				4. TEMPO DE RETORNO=				ME	11. TEMPO DE RETORNO=				12. TEMPO DE RETORNO=					
TUDO	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	
N	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
LN	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
LN3	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
EXP	1.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
*EXP	1.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
GAM	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
P3	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
LP3	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
GH	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
GEV	0.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
*AK	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

ORS: EM 100% DOS GRUPOS O MÉTODO MAK CONVERGIU

Tabela 4.8

VALORES DE PERIGO PARA O RISCO DE COLATINA OBTIDAS COM 1 PICOS POR ANO

N= 20 LG= 500

TEMPO DE RETORNO= 2. POPULACAO= 3745.0									TEMPO DE RETORNO= 10. POPULACAO= 5869.0								
MF	TEPO	EPD	EMA		<0.9	0.9-1.0	1.0-1.1	>1.1	MED	EMA		<0.9	0.9-1.0	1.0-1.1	>1.1		
A	1.16	0.00	0.16	0.00	0.007	0.038	0.222	0.738	0.98	0.00	0.09	0.00	0.246	0.334	0.274	0.146	
LA	1.12	0.00	0.13	0.00	0.006	0.050	0.342	0.592	0.96	0.00	0.09	0.00	0.296	0.352	0.260	0.092	
LN3	1.12	0.00	0.13	0.00	0.006	0.055	0.344	0.564	0.99	0.00	0.09	0.00	0.236	0.322	0.286	0.156	
FXD	1.06	0.00	0.08	0.00	0.024	0.172	0.496	0.308	0.99	0.00	0.09	0.00	0.236	0.326	0.284	0.154	
*FXD	1.06	0.00	0.08	0.00	0.016	0.180	0.504	0.300	0.99	0.00	0.09	0.00	0.162	0.348	0.304	0.166	
GAX	1.13	0.00	0.13	0.00	0.004	0.054	0.322	0.620	0.99	0.01	0.09	0.00	0.226	0.318	0.294	0.162	
P3	1.09	0.00	0.10	0.00	0.024	0.129	0.408	0.440	0.98	0.00	0.09	0.00	0.240	0.336	0.280	0.144	
LP3	1.09	0.00	0.10	0.00	0.018	0.118	0.438	0.426	0.97	0.01	0.09	0.00	0.260	0.340	0.262	0.118	
GU	1.11	0.00	0.11	0.00	0.017	0.086	0.392	0.510	0.99	0.00	0.09	0.00	0.236	0.326	0.282	0.156	
GFV	1.10	0.00	0.11	0.00	0.014	0.096	0.396	0.494	0.98	0.00	0.09	0.00	0.256	0.332	0.272	0.140	
GAX	1.07	0.00	0.09	0.00	0.026	0.156	0.454	0.330	0.99	0.00	0.09	0.00	0.228	0.342	0.278	0.152	

TEMPO DE RETORNO= 25. POPULACAO= 6944.0									TEMPO DE RETORNO= 50. POPULACAO= 7812.0								
MF	TEPO	EPD	EMA		<0.9	0.9-1.0	1.0-1.1	>1.1	MED	EMA		<0.9	0.9-1.0	1.0-1.1	>1.1		
A	0.90	0.00	0.13	0.00	0.532	0.290	0.136	0.042	0.85	0.00	0.16	0.00	0.684	0.212	0.088	0.016	
LA	0.90	0.00	0.12	0.00	0.510	0.298	0.146	0.046	0.87	0.00	0.15	0.00	0.626	0.254	0.094	0.026	
LN3	0.93	0.01	0.12	0.00	0.422	0.264	0.200	0.094	0.90	0.00	0.14	0.00	0.556	0.214	0.140	0.090	
FXD	0.97	0.01	0.11	0.00	0.294	0.300	0.238	0.168	0.97	0.01	0.12	0.00	0.320	0.272	0.222	0.186	
*FXD	0.94	0.01	0.10	0.00	0.254	0.306	0.256	0.184	0.99	0.01	0.11	0.00	0.268	0.296	0.246	0.190	
GAX	0.93	0.01	0.12	0.00	0.444	0.268	0.174	0.094	0.89	0.01	0.14	0.00	0.570	0.240	0.122	0.068	
P3	0.95	0.01	0.12	0.00	0.348	0.304	0.206	0.142	0.94	0.01	0.14	0.00	0.420	0.254	0.174	0.152	
LP3	0.95	0.01	0.12	0.00	0.342	0.296	0.224	0.138	0.95	0.01	0.14	0.01	0.368	0.252	0.194	0.186	
GU	0.95	0.01	0.11	0.00	0.362	0.320	0.260	0.116	0.93	0.01	0.13	0.00	0.460	0.258	0.176	0.106	
GFV	0.94	0.01	0.11	0.00	0.376	0.306	0.196	0.122	0.93	0.01	0.13	0.00	0.448	0.254	0.180	0.118	
GAX	1.00	0.01	0.11	0.00	0.254	0.304	0.224	0.218	1.02	0.01	0.13	0.00	0.244	0.246	0.206	0.304	

Q95: EM 1003 DOS GRUPOS O METODO WAK CONVERGIU

Tabela 4.9

102.

1. TEMPO DE RETORNO=		2. POPULACAO= 3745.0		3. TEMPO DE RETORNO=		4. POPULACAO= 5669.0										
ME	100.0	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	MED	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1			
K	1.15	0.00	0.15	0.00	0.000	0.034	0.272	0.694	0.92	0.00	0.10	0.00	0.424	0.386	0.164	0.026
LV	1.10	0.00	0.10	0.00	0.010	0.102	0.432	0.458	0.91	0.00	0.10	0.00	0.462	0.380	0.138	0.020
LV3	1.11	0.00	0.12	0.00	0.008	0.076	0.358	0.558	0.96	0.00	0.09	0.00	0.314	0.372	0.224	0.090
FYP	1.06	0.00	0.09	0.00	0.015	0.178	0.470	0.316	0.99	0.00	0.09	0.00	0.202	0.358	0.288	0.152
*EXP	1.06	0.00	0.08	0.00	0.016	0.164	0.510	0.310	0.99	0.00	0.08	0.00	0.182	0.390	0.298	0.130
GAK	1.13	0.00	0.13	0.00	0.006	0.048	0.310	0.636	0.95	0.00	0.09	0.00	0.328	0.366	0.226	0.080
P3	1.07	0.00	0.09	0.00	0.024	0.146	0.462	0.368	0.98	0.00	0.09	0.00	0.254	0.344	0.266	0.136
LP3	1.06	0.00	0.08	0.00	0.022	0.174	0.508	0.296	0.97	0.00	0.08	0.00	0.274	0.360	0.266	0.100
GB	1.11	0.00	0.11	0.00	0.012	0.084	0.396	0.508	0.96	0.00	0.09	0.00	0.292	0.378	0.228	0.102
GFV	1.09	0.00	0.10	0.00	0.016	0.112	0.464	0.408	0.96	0.00	0.09	0.00	0.304	0.374	0.234	0.088
WAK	1.06	0.00	0.08	0.00	0.026	0.202	0.486	0.286	0.99	0.00	0.09	0.00	0.192	0.362	0.286	0.160

! TEMPO DE RETORNO= 25.					POPULACAO= 6994.0				! TEMPO DE RETORNO= 50.					POPULACAO= 7812.0			
ME	MED		EVA		<0.9	0.9-1.0	1.0-1.1	>1.1	ME	MED		EVA		<0.9	0.9-1.0	1.0-1.1	>1.1
N	0.94	0.00	0.16	0.00	0.750	0.210	0.038	0.002	0.79	0.00	0.21	0.00	0.874	0.114	0.012	0.000	
LN	0.95	0.00	0.15	0.00	0.707	0.232	0.064	0.002	0.83	0.00	0.18	0.00	0.772	0.180	0.048	0.000	
LN3	0.91	0.01	0.12	0.00	0.520	0.268	0.148	0.064	0.89	0.01	0.15	0.00	0.586	0.238	0.114	0.062	
FYP	0.98	0.01	0.10	0.00	0.268	0.316	0.266	0.150	0.97	0.01	0.10	0.00	0.300	0.292	0.246	0.162	
*EXP	0.97	0.00	0.09	0.00	0.247	0.384	0.246	0.128	0.97	0.00	0.09	0.00	0.276	0.350	0.242	0.132	
GAV	0.99	0.00	0.13	0.00	0.542	0.298	0.126	0.034	0.86	0.01	0.15	0.00	0.666	0.220	0.092	0.022	
P3	0.96	0.01	0.11	0.00	0.330	0.314	0.218	0.138	0.95	0.01	0.13	0.00	0.388	0.270	0.186	0.156	
LP3	0.97	0.01	0.11	0.00	0.302	0.292	0.240	0.166	1.00	0.01	0.13	0.00	0.286	0.256	0.212	0.246	
GU	0.93	0.00	0.11	0.00	0.430	0.318	0.202	0.050	0.91	0.01	0.12	0.00	0.498	0.310	0.144	0.048	
GFV	0.94	0.01	0.11	0.00	0.386	0.322	0.192	0.100	0.93	0.01	0.12	0.00	0.422	0.272	0.192	0.114	
*AF	1.00	0.01	0.11	0.00	0.242	0.284	0.245	0.228	1.02	0.01	0.13	0.00	0.244	0.238	0.210	0.308	

ORS: Fu 100% DOS GRUPOS O METODO WAK CONVERGIO

Tabela 4.10

VALORES DOS PROJETOS PARA O PISIO DE COLATINA OBTIDAS COM 3 PICOS POR ANO

N= 20 PR= 3.00

TEMPO DE RETORNO= 2. POPULACAO= 3745.0															
TEMPO DE RETORNO= 10. POPULACAO= 5869.0															
ME	TEPO	REC	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	REC	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1		
N	1.12	0.00	0.13	0.00	0.007	0.052	0.352	0.88	0.00	0.12	0.00	0.605	0.318	0.073	0.005
LN	1.07	0.00	0.08	0.00	0.011	0.139	0.552	0.88	0.00	0.13	0.00	0.627	0.305	0.064	0.005
IN3	1.10	0.00	0.11	0.00	0.009	0.085	0.430	0.94	0.00	0.09	0.00	0.361	0.366	0.214	0.054
FYP	1.07	0.00	0.08	0.00	0.011	0.148	0.491	0.99	0.00	0.08	0.00	0.191	0.380	0.298	0.132
*FYP	1.06	0.00	0.07	0.00	0.018	0.180	0.534	0.96	0.00	0.08	0.00	0.227	0.473	0.248	0.052
GAY	1.12	0.00	0.12	0.00	0.009	0.055	0.386	0.93	0.00	0.10	0.00	0.409	0.384	0.170	0.036
P3	1.07	0.00	0.08	0.00	0.016	0.145	0.498	0.97	0.00	0.08	0.00	0.250	0.368	0.257	0.125
LP3	1.05	0.00	0.07	0.00	0.016	0.216	0.550	0.97	0.00	0.08	0.00	0.239	0.362	0.277	0.102
CU	1.10	0.00	0.10	0.00	0.011	0.082	0.436	0.95	0.00	0.09	0.00	0.325	0.398	0.227	0.050
GFV	1.07	0.00	0.09	0.00	0.011	0.127	0.511	0.95	0.00	0.09	0.00	0.309	0.384	0.243	0.064
WAK	1.06	0.00	0.08	0.00	0.018	0.182	0.514	1.00	0.00	0.08	0.00	0.150	0.375	0.307	0.168

TEMPO DE RETORNO= 25. POPULACAO= 6994.0															
TEMPO DE RETORNO= 50. POPULACAO= 7812.0															
ME	TEPO	REC	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1	REC	EMA	<0.9	0.9-1.0	1.0-1.1	>1.1		
N	0.80	0.00	0.20	0.00	0.893	0.098	0.009	0.75	0.00	0.25	0.00	0.977	0.023	0.000	0.000
LN	0.83	0.00	0.17	0.00	0.811	0.166	0.023	0.80	0.00	0.20	0.00	0.870	0.114	0.016	0.000
LN3	0.90	0.01	0.13	0.00	0.552	0.277	0.125	0.88	0.01	0.15	0.00	0.611	0.232	0.107	0.050
FYP	0.97	0.01	0.09	0.00	0.286	0.364	0.245	0.97	0.01	0.09	0.00	0.289	0.345	0.239	0.127
*FYP	0.94	0.00	0.09	0.00	0.336	0.416	0.207	0.93	0.00	0.09	0.00	0.377	0.389	0.193	0.041
GAY	0.87	0.00	0.14	0.00	0.650	0.241	0.095	0.84	0.00	0.17	0.00	0.734	0.195	0.059	0.011
P3	0.96	0.01	0.11	0.00	0.319	0.320	0.225	0.96	0.01	0.12	0.00	0.361	0.280	0.205	0.155
LP3	1.00	0.01	0.10	0.00	0.239	0.293	0.259	1.03	0.01	0.13	0.00	0.205	0.250	0.214	0.332
CU	0.91	0.00	0.11	0.00	0.468	0.357	0.150	0.89	0.00	0.13	0.00	0.559	0.298	0.123	0.020
GFV	0.94	0.01	0.10	0.00	0.382	0.332	0.205	0.94	0.01	0.12	0.00	0.395	0.298	0.191	0.116
WAK	1.00	0.01	0.11	0.00	0.225	0.300	0.234	1.02	0.01	0.14	0.00	0.245	0.250	0.198	0.307

ORR: EM 100% DOS GRUPOS O METODO WAK CONVERTEU

Tabela 4.11

As tabelas 4.12 e 4.13 listam, para cada tempo de recorrência e tamanho da amostra, as distribuições que apresentaram o menor erro médio absoluto na estimação dos quantis, bem como os valores destes erros, tanto no caso das séries de máximos anuais, como nas séries parciais. É possível verificar-se, inicialmente, que para as séries parciais a distribuição exponencial foi a que apresentou os menores erros nas estimativas de vazões extremas, ao passo que para as séries de máximos anuais a distribuição Gumbel foi a que apresentou melhor desempenho. Este resultado confirma aquele obtido por Damazio et al. (1983), segundo o qual as distribuições de dois parâmetros seriam mais eficientes para estimar vazões com tempos de recorrência muito elevados. Para os tempos de recorrência mais baixos, várias distribuições de probabilidades apresentam comportamento semelhante, tanto no caso de séries parciais, quanto no de máximos anuais.

Uma objeção inicial pode ser levantada a este resultado, uma vez que o modelo XADREZ, pela sua própria formulação, poderia estar gerando máximos anuais com distribuição Gumbel e séries parciais com distribuição exponencial, o que seria até explicável mediante certas hipóteses já discutidas aqui (item 2.2.1). Buscando dirimir esta dúvida é que foram geradas séries sintéticas com o modelo DIANA, cujos resultados estão apresentados no próximo item.

Com relação à magnitude dos erros absolutos cometidos, pode-se notar que os métodos de séries parciais (exceto quando aplicados às séries SPL) forneceram resultados melhores que os métodos de máximos anuais. Em geral, estes erros decrescem à medida em que a taxa anual de picos é maior, ou que o tamanho da

SÉRIE N	SMA				SP1				SP2				SP3			
	2	10	25	50	2	10	25	50	2	10	25	50	2	10	25	50
5	(1) 0,13	(1) 0,17	(1) 0,20	GU 0,21	LN 0,12	LN/EXP* 0,17	EXP*/WAK 0,20	EXP* 0,22	(1) 0,12	EXP* 0,14	EXP* 0,16	EXP* 0,16	(1) 0,12	EXP*/WAK 0,14	EXP* 0,15	EXP* 0,15
10	(1) 0,09	(1) 0,12	(1) 0,14	GU 0,15	(1) 0,08	(1) 0,12	EXP*/WAK 0,14	EXP* 0,15	(1) 0,07	EXP*/WAK 0,11	EXP* 0,12	EXP* 0,12	LN 0,06	EXP* 0,10	EXP* 0,11	EXP* 0,11
15	GU/LN 0,07	(1) 0,10	GU 0,11	GU 0,12	(1) 0,07	(1) 0,10	EXP*/WAK 0,12	EXP* 0,13	(1) 0,06	EXP* 0,09	EXP* 0,10	EXP* 0,10	LN 0,05	(1) 0,09	EXP*/WAK 0,10	EXP* 0,10
20	GU 0,06	(1) 0,09	(1) 0,10	GU/EXP 0,11	(1) 0,06	(1) 0,09	EXP* 0,10	EXP* 0,11	(1) 0,05	EXP* 0,08	EXP* 0,09	EXP* 0,09	(1) 0,05	(1) 0,08	EXP*/EXP 0,09	EXP*/EXP 0,09
25	(1) 0,06	(1) 0,08	(1) 0,09	GU 0,09	LN 0,05	(1) 0,08	EXP* 0,09	EXP* 0,10	(1) 0,05	EXP* 0,07	EXP* 0,08	EXP* 0,08	LN 0,04	EXP*/WAK 0,07	EXP 0,08	EXP*/EXP* 0,09
30	GU/LN 0,05	(1) 0,07	(1) 0,08	GU 0,08	(1) 0,05	(1) 0,07	EXP* 0,08	EXP* 0,08	LN 0,04	(1) 0,07	EXP* 0,07	EXP*/EXP 0,08	(1) 0,04	EXP 0,06	EXP 0,07	EXP 0,08
40	(1) 0,05	LN/GEV 0,05	LN 0,06	LN/GU 0,07	(1) 0,04	EXP* 0,05	EXP* 0,06	EXP* 0,07	(1) 0,04	EXP* 0,05	EXP*/EXP 0,06	EXP* 0,06	LN 0,03	EXP*/WAK 0,05	EXP*/LP3 0,06	EXP 0,06
50	(1) 0,04	(1) 0,05	(1) 0,06	LN/GU 0,06	(1) 0,04	(1) 0,05	EXP*/EXP* 0,06	EXP* 0,06	(1) 0,03	(1) 0,05	EXP* 0,05	EXP*/EXP 0,06	(1) 0,03	(1) 0,05	EXP 0,05	EXP 0,06
60	(1) 0,04	(1) 0,05	(1) 0,06	LN/GU 0,06	LN 0,03	(1) 0,05	EXP* 0,05	EXP* 0,06	(1) 0,03	(1) 0,05	EXP* 0,05	EXP*/EXP 0,06	(1) 0,03	(1) 0,05	(1) 0,06	EXP 0,06

Tabela 4.12 - Erros Méricos Absolutos das Distribuições de Melhor Desempenho. - Modelo Xadrez, Colatina
Pequenos Tempos de Recorrência

1 - Várias distribuições apresentaram o mesmo desempenho.

(*) - Método da Máxima Verossimilhança.

SÉRIE T N	SMA				SP1				SP2				SP3			
	100	500	1.000	10.000	100	500	1.000	10.000	100	500	1.000	10.000	100	500	1.000	10.000
5	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*
	0,22	0,24	0,25	0,26	0,23	0,26	0,26	0,30	0,17	0,18	0,18	0,19	0,16	0,16	0,16	0,15
10	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*
	0,15	0,17	0,17	0,19	0,16	0,17	0,18	0,20	0,12	0,12	0,12	0,13	0,12	0,12	0,12	0,11
15	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*
	0,13	0,14	0,14	0,16	0,13	0,15	0,15	0,17	0,10	0,11	0,11	0,12	0,11	0,11	0,11	0,10
20	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*
	0,11	0,12	0,13	0,14	0,11	0,13	0,13	0,15	0,09	0,10	0,10	0,11	0,10	0,10	0,10	0,09
25	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP	EXP	EXP	EXP*
	0,10	0,11	0,11	0,12	0,10	0,11	0,12	0,14	0,08	0,09	0,09	0,10	0,09	0,10	0,10	0,09
30	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP	EXP	EXP	EXP*
	0,09	0,10	0,10	0,11	0,09	0,10	0,10	0,12	0,08	0,08	0,09	0,09	0,08	0,08	0,09	0,08
40	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP	EXP	EXP	EXP*
	0,07	0,08	0,08	0,09	0,07	0,08	0,08	0,10	0,06	0,07	0,07	0,08	0,07	0,07	0,07	0,07
50	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP	EXP	EXP	EXP*
	0,07	0,07	0,07	0,08	0,06	0,07	0,07	0,09	0,06	0,06	0,06	0,07	0,06	0,06	0,06	0,06
60	GU	GU	GU	GU	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP*	EXP	EXP	EXP	EXP
	0,06	0,07	0,07	0,08	0,06	0,07	0,07	0,08	0,06	0,06	0,06	0,07	0,06	0,06	0,06	0,07

Tabela 4.13 - Erros Médios Absolutos das Distribuições de Melhor Desempenho.
Grandes Tempos de Recorrência - Modelo Xadrez, Colatina

(*) - Método da Máxima Verossimilhança.

amostra é maior. Esta variação pode ser vista através das figuras 4.4 a 4.6, que apresentam curvas com os erros médios nas abscissas e o tamanho das amostras nas ordenadas para três diferentes tempos de recorrência: 100, 1000 e 10000 anos. Observa-se que o aumento do tamanho das amostras diminui o erro médio das estimativas, como era de se esperar, diminuindo também a diferença entre os desempenhos dos métodos de séries parciais para os métodos de máximos anuais. Estas figuras sugerem também que, para pequenas amostras (n menor que 15), o aumento de precisão quando se acrescentam mais 5 anos de registro é significativo, o mesmo não acontecendo para amostras maiores que 25 anos.

As figuras 4.7 a 4.9 mostram a variação da probabilidade de subdimensionamento com o aumento do tamanho das amostras para os mesmos tempos de recorrência já vistos. Neste caso, como pode ser visto, 5 anos a mais de registro podem ser muito valiosos, pois a probabilidade de subdimensionamento diminui bastante, principalmente na estimação de vazões decamilenares. Nota-se que, neste caso, novamente o método de séries parciais apresentou desempenho mais satisfatório que o método de máximos anuais.

Deve-se ressaltar, aqui, que a objeção levantada anteriormente à escolha das distribuições com melhores desempenhos não mais se aplica. Isto porque, mesmo que o modelo XADREZ fornecesse, por construção, séries parciais exponencialmente distribuidas (e, conseqüentemente, máximos anuais com distribuição Gumbel), o que importa agora é qual dos dois métodos analisados apresenta os menores erros médios absolutos, na estimação dos quantis.

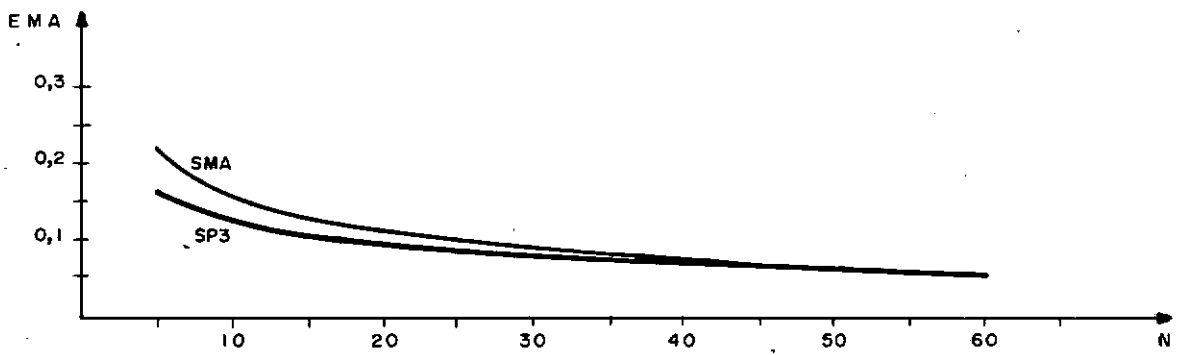


FIGURA 4.4 — TEMPO DE RECORRÊNCIA = 100 ANOS.

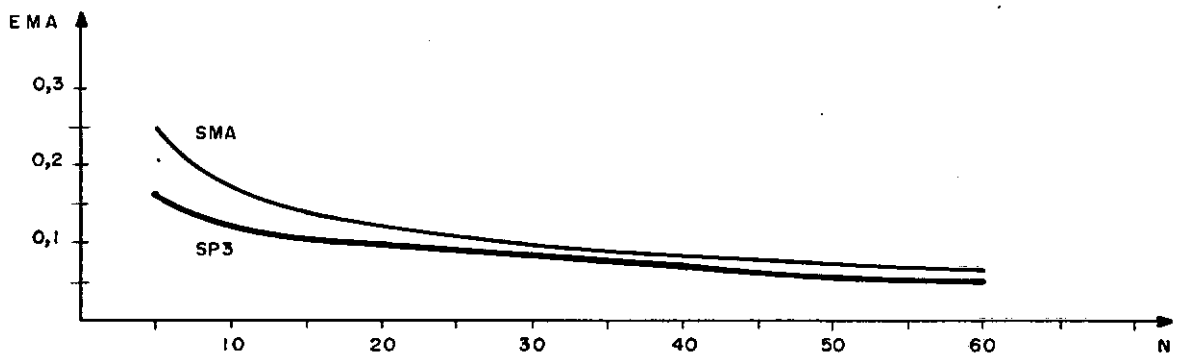


FIGURA 4.5 — TEMPO DE RECORRÊNCIA = 1000 ANOS.

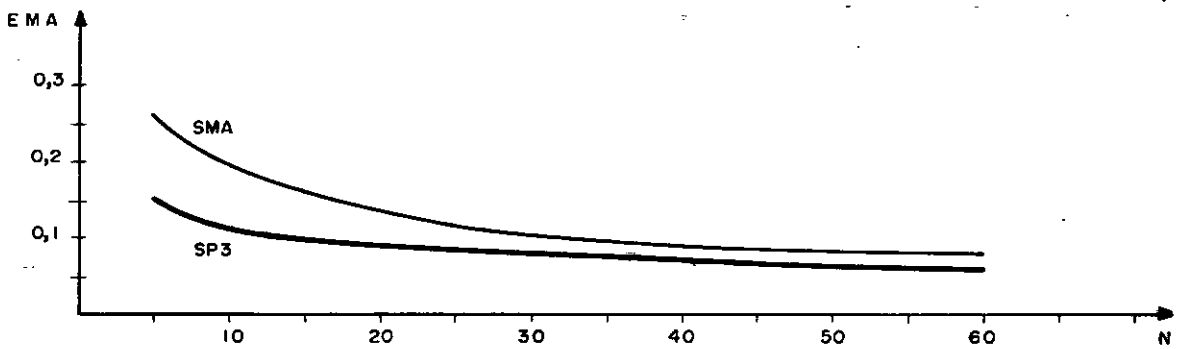


FIGURA 4.6 — TEMPO DE RECORRÊNCIA = 10000 ANOS.

EFEITO DO TAMANHO DA AMOSTRA SOBRE O ERRO MÉDIO ABSOLUTO.

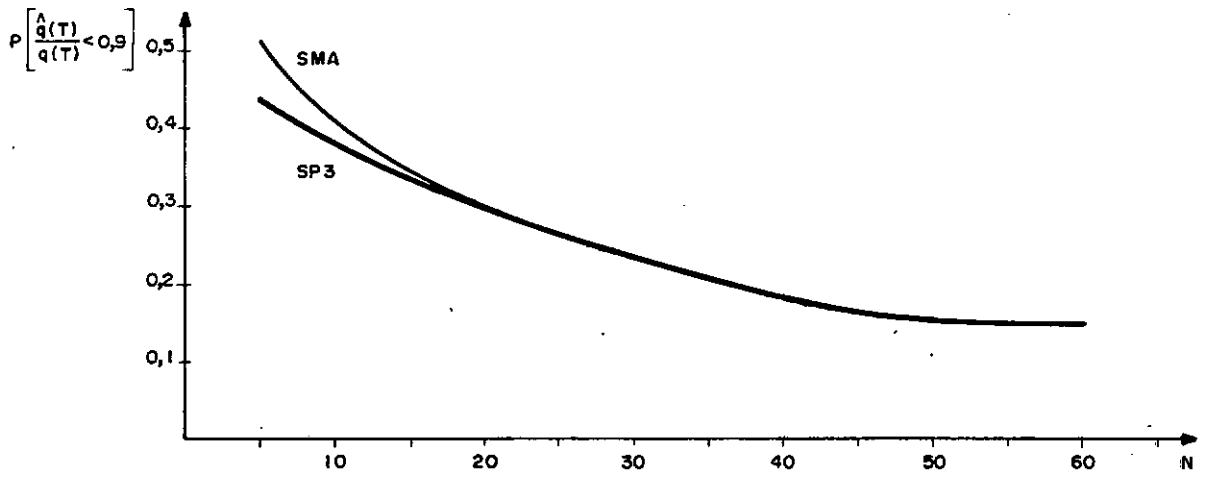


FIGURA 4.7 — TEMPO DE RECORRÊNCIA = 100 ANOS.

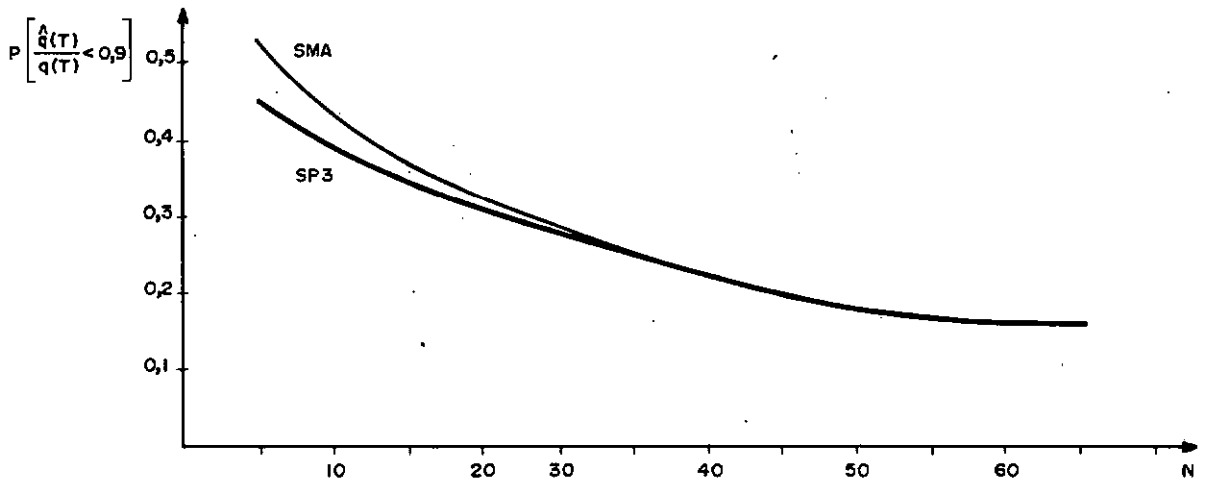


FIGURA 4.8 — TEMPO DE RECORRÊNCIA = 1000 ANOS.

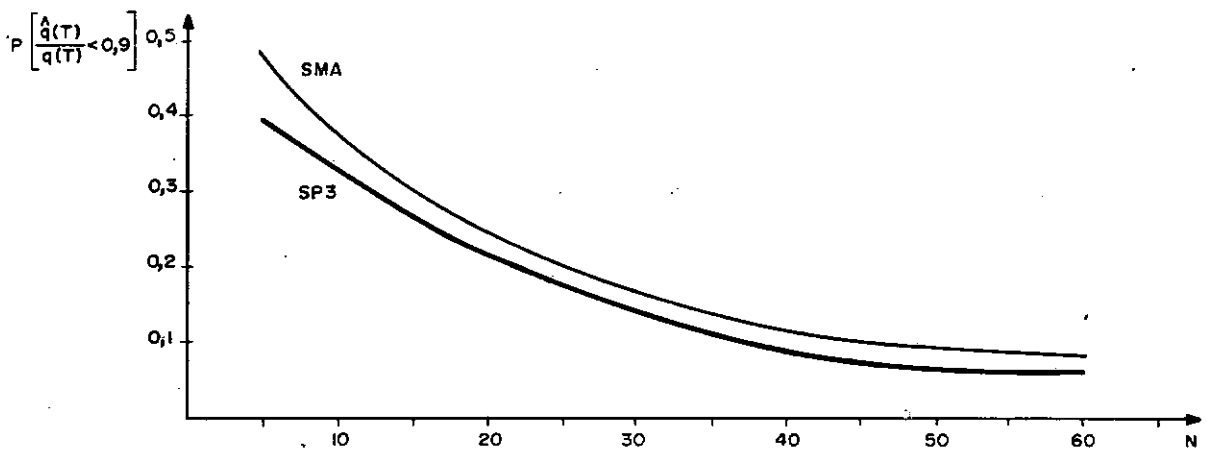


FIGURA 4.9 — TEMPO DE RECORRÊNCIA = 10000 ANOS.

EFEITO DO TAMANHO DA AMOSTRA SOBRE A PROBABILIDADE DE SUB-ESTIMATIVAS.

Todavia, para populações diferentes este resultado não se repete necessariamente. Isto é o que se pretende verificar no próximo item.

4.4 - Aplicação a Séries Sintéticas - Modelo DIANA

Os resultados obtidos na experiência apresentada no item 4.3.2 podem ser questionados uma vez que o modelo XADREZ, pela sua estrutura, poderia estar conduzindo àquelas conclusões. Na tentativa de dirimir estas dúvidas, foram geradas novas séries sintéticas com o modelo DIANA (Kelman et al., 1983). Como este modelo permite a geração de várias séries multivariadamente, foi tentada uma aplicação de cunho regional. Para isto, seria ideal a geração multivariada de séries para todos os postos de uma região, o que seria inviável do ponto de vista computacional. Assim foram escolhidos quatro postos fluviométricos da bacia do rio Doce, para aplicação do modelo DIANA. Estes postos estão listados na tabela 4.14. Foram geradas 10.000 séries de vazões diárias afluentes a estes 4 postos, cada série contendo 212 dias, correspondentes ao período chuvoso da bacia.

A comparação dos índices das séries geradas com os das séries históricas revelou que apenas em dois casos (rio Doce, em Colatina, e Rio Piranga, em Porto Firme) a calibração do modelo foi satisfatória. Uma vez que a comparação entre os métodos anuais e os métodos de séries parciais exige, a princípio, apenas séries completas de vazões diárias, não foram feitas maiores tentativas de se obterem calibrações melhores, mesmo sabendo ser o modelo utilizado suficientemente versátil para fornecê-las. Além disso, não foi possível manter, nas séries geradas, a mesma estru

tura de correlação espacial observada nas séries históricas, o que impediu a realização de experiências com metodologias de regionalização.

CÓDIGO DNAEE	R I O	LOCALIDADE	ÁREA Km ²
56075000	Piranga	Porto Firme	4171
56720000	Doce	Cachoeira Escura	24245
56990000	Manhuaçu	São Sebastião da Encruzilhada	8454
56994500	Doce	Colatina	76615

Tabela 4.14 - Postos da bacia do rio Doce escolhidos para aplicação do modelo DIANA.

Entretanto, de posse de quatro séries de 1000 anos de vazões médias diárias, foi possível a aplicação da metodologia de avaliação descrita no item 4.2.

A tabela 4.15 apresenta algumas características das quatro populações de máximos anuais geradas: os quatro primeiros momentos centrais e os valores de população para tempos de recorrência de 100, 250, 500 e 1000 anos.

O experimento realizado com as séries geradas pelo modelo XADREZ foi repetido para estas quatro séries, segundo a mesma metodologia. Entretanto, duas limitações se impuseram:

					TEMPO DE RECORRÊNCIA			
POSTO	\bar{Q}	CV	ASS	CUR	100	250	500	1000
56075000	434	0,407	1,342	5,895	1028	1125	1215	1292
56720000	1935	0,452	1,607	7,232	4870	5531	6128	6771
56990000	585	0,531	2,219	11,911	1744	2064	2314	2720
56994500	4384	0,377	1,125	4,864	9682	10761	11418	11966

Tabela 4.15 - Estatísticas dos máximos anuais gerados pelo modelo DIANA.

- 1) o maior tempo de recorrência estudado foi 1000 anos, pois com 10000 séries não seria possível determinar com precisão quantis mais elevados e
- 2) a precisão das estimativas dos índices de avaliação foi menor, pois só estavam disponíveis 1000 séries completas.

As tabelas 4.16 a 4.19 mostram um resumo dos resultados obtidos a partir dos 4 cenários. Em três deles (Colatina, Cachoeira Escura e Porto Firme), estes resultados confirmaram os da experiência anterior. Os métodos de séries parciais apresentaram desempenhos melhores que os métodos de máximos anuais para a estimação de quantis de grandes tempos de recorrência a partir de amostra pequenas, principalmente. Entretanto, a diferença entre os dois métodos já não foi tão significativa.

Para o posto de S, Sebastião da Encruzilhada, a si

SÉRIE T N	SMA				SP1				SP2				SP3			
	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000
5	EXP 0,27	GU 0,28	GU 0,29	GU 0,29	EXP* 0,26	EXP 0,28	EXP 0,29	EXP 0,30	EXP* 0,21	EXP* 0,22	EXP* 0,23	EXP* 0,24	EXP* 0,18	EXP* 0,18	EXP* 0,19	EXP* 0,19
10	EXP 0,22	GU 0,22	GU 0,23	GU 0,24	EXP* 0,18	EXP* 0,19	EXP* 0,20	EXP* 0,21	EXP* 0,15	EXP* 0,15	EXP* 0,16	EXP* 0,16	EXP* 0,13	EXP* 0,13	EXP* 0,13	EXP* 0,13
15	GU 0,17	GU 0,18	GU 0,18	GU 0,19	EXP* 0,12	EXP* 0,14	EXP* 0,15	EXP* 0,16	EXP* 0,11	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,11	EXP* 0,10	EXP* 0,11	EXP* 0,11
20	GU 0,14	GU 0,15	GU 0,15	GU 0,16	EXP* 0,10	EXP* 0,11	EXP* 0,12	EXP* 0,13	EXP* 0,10	EXP* 0,11	EXP* 0,11	EXP* 0,11	EXP* 0,10	EXP* 0,09	EXP* 0,09	EXP* 0,09
25	GU 0,13	GU 0,13	GU 0,13	GU 0,14	EXP* 0,10	EXP* 0,11	EXP* 0,12	EXP* 0,13	EXP* 0,09	EXP* 0,10	EXP* 0,10	EXP* 0,11	EXP* 0,09	EXP* 0,08	EXP* 0,08	EXP* 0,08
30	GU 0,11	GU 0,11	GU 0,12	GU 0,12	EXP* 0,09	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,08	EXP* 0,08	EXP* 0,09	EXP* 0,09	EXP 0,09	EXP* 0,07	EXP* 0,07	EXP* 0,07
40	EXP 0,12	GU 0,12	GU 0,12	GU 0,12	EXP* 0,08	EXP* 0,10	EXP* 0,10	EXP* 0,11	EXP* 0,07	EXP* 0,08	EXP* 0,08	EXP* 0,09	EXP* 0,08	EXP* 0,06	EXP* 0,06	EXP* 0,06
50	EXP 0,08	LN 0,07	GU 0,07	GU 0,08	GU 0,08	GU 0,08	GU 0,09	GU 0,09	EXP* 0,06	EXP* 0,06	EXP* 0,07	EXP* 0,07	EXP 0,06	EXP* 0,05	EXP* 0,05	EXP* 0,05
60	EXP 0,09	GU 0,09	GU 0,09	GU 0,09	GU 0,06	GU 0,07	GU 0,08	GU 0,08	EXP* 0,06	EXP* 0,06	EXP* 0,06	EXP* 0,07	EXP 0,06	EXP* 0,05	EXP* 0,05	EXP* 0,04

Tabela 4.16 - Erros Médios Absolutos das Distribuições de Melhor Desempenho. - Modelo Diana, Porto Firme

(*) - Método de Máxima Verossimilhança.

SERIE		SMA				SP1				SP2				SP3			
N	T	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000
		EXP 0,28	GU 0,29	GU 0,29	EXP 0,31	EXP* 0,27	EXP* 0,28	EXP* 0,29	EXP* 0,29	EXP* 0,22	EXP* 0,22	EXP* 0,23	EXP* 0,23	EXP* 0,18	EXP* 0,18	EXP* 0,19	EXP* 0,20
10		EXP 0,22	EXP 0,23	EXP 0,24	EXP 0,25	EXP* 0,16	EXP* 0,17	EXP* 0,18	EXP* 0,18	EXP* 0,15	EXP* 0,15	EXP* 0,16	EXP* 0,17	EXP* 0,15	EXP* 0,15	EXP* 0,16	EXP* 0,17
15		GU 0,16	GU 0,17	GU 0,18	EXP 0,19	EXP* 0,14	EXP* 0,14	EXP* 0,15	EXP* 0,15	EXP* 0,12	EXP* 0,12	EXP* 0,13	EXP* 0,14	EXP 0,12	EXP 0,13	EXP 0,13	EXP 0,14
20		LN 0,15	LN 0,16	EXP 0,17	EXP 0,18	EXP* 0,12	EXP* 0,12	EXP* 0,13	EXP* 0,13	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,13	EXP 0,11	EXP 0,11	EXP 0,11	EXP 0,12
25		EXP 0,14	GU 0,14	GU 0,15	EXP 0,16	EXP* 0,10	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,09	EXP* 0,10	EXP* 0,11	EXP* 0,12	EXP 0,10	EXP 0,10	EXP 0,11	EXP 0,12
30		LN 0,12	LN 0,12	LN 0,13	LN 0,14	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,12	EXP 0,09	EXP 0,09	EXP* 0,11	EXP* 0,12	EXP 0,10	EXP 0,10	EXP 0,11	EXP 0,12
40		GU 0,13	GU 0,13	GU 0,14	EXP 0,15	EXP* 0,09	EXP* 0,10	EXP* 0,10	EXP* 0,10	EXP 0,09	EXP 0,09	EXP 0,10	EXP 0,10	EXP 0,09	EXP 0,09	EXP 0,10	EXP 0,10
50		GU 0,11	GU 0,11	GU 0,12	GU 0,13	EXP 0,07	EXP 0,07	EXP 0,07	EXP 0,08	EXP 0,06	EXP 0,06	EXP 0,07	EXP 0,08	EXP 0,07	EXP 0,07	EXP 0,08	EXP 0,09
60		GU 0,09	GU 0,09	LN 0,10	EXP 0,11	EXP* 0,08	EXP* 0,08	EXP* 0,09	EXP* 0,10	EXP 0,07	EXP 0,08	EXP 0,08	EXP 0,09	EXP 0,07	EXP 0,08	EXP 0,08	EXP 0,10

Tabela 4.17 - Erros Médios Absolutos das Distribuições de Melhor Desempenho. - Modelo Diana, Cachoeira Escura

(*) - Método de Máxima Verossimilhança.

SÉRIE T N	SMA				SP1				SP2				SP3			
	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000
5	EXP* 0,29	EXP* 0,30	EXP* 0,30	EXP* 0,31	EXP* 0,28	EXP* 0,29	EXP* 0,30	EXP* 0,33	EXP* 0,26	EXP* 0,28	EXP* 0,29	EXP* 0,33	EXP 0,28	EXP 0,30	EXP 0,31	EXP 0,35
10	EXP* 0,20	EXP* 0,21	EXP* 0,21	EXP* 0,22	EXP* 0,21	EXP* 0,22	EXP* 0,23	EXP* 0,27	EXP 0,23	EXP 0,25	EXP 0,26	EXP 0,30	EXP 0,24	EXP 0,26	EXP 0,27	EXP 0,31
15	EXP* 0,17	EXP* 0,18	EXP* 0,18	EXP* 0,17	EXP* 0,17	EXP* 0,19	EXP* 0,20	EXP* 0,25	EXP 0,21	EXP 0,23	EXP 0,24	EXP 0,28	EXP 0,21	EXP 0,23	EXP 0,25	EXP 0,29
20	EXP* 0,15	EXP* 0,15	EXP* 0,15	EXP* 0,14	EXP* 0,17	EXP* 0,18	EXP* 0,20	EXP* 0,25	EXP 0,19	EXP 0,21	EXP 0,23	EXP 0,27	EXP 0,20	EXP 0,22	GEV 0,24	GEV 0,27
25	EXP* 0,14	EXP* 0,14	EXP* 0,14	EXP* 0,13	EXP* 0,14	EXP* 0,17	EXP* 0,18	EXP* 0,23	EXP 0,17	EXP 0,19	EXP 0,21	EXP 0,26	EXP 0,18	GEV 0,21	GEV 0,22	GEV 0,25
30	EXP* 0,13	EXP* 0,13	EXP* 0,13	EXP* 0,12	EXP* 0,13	EXP* 0,16	EXP* 0,17	EXP* 0,24	LP 3 0,16	EXP 0,18	EXP 0,20	EXP 0,25	LP 3 0,17	GEV 0,20	GEV 0,21	GEV 0,24
40	EXP* 0,11	EXP* 0,11	EXP* 0,11	EXP* 0,09	EXP* 0,13	EXP* 0,16	EXP* 0,18	EXP* 0,23	LP 3 0,13	EXP 0,17	EXP 0,19	EXP 0,24	LP 3 0,15	GEV 0,19	GEV 0,20	GEV 0,23
50	EXP* 0,12	EXP* 0,12	EXP* 0,11	EXP* 0,09	EXP* 0,14	EXP* 0,16	EXP* 0,18	EXP* 0,24	LP 3 0,12	EXP 0,18	EXP 0,20	GEV 0,25	LP 3 0,15	GEV 0,18	GEV 0,19	GEV 0,22
60	EXP* 0,10	EXP* 0,10	EXP* 0,10	EXP* 0,08	EXP* 0,13	EXP 0,16	EXP 0,18	EXP 0,22	LP 3 0,11	LP 3 0,16	EXP 0,18	GEV 0,22	LP 3 0,12	GEV 0,16	GEV 0,16	GEV 0,19

Tabela 4.18 - Erros Médios Absolutos das Distribuições de Melhor Desempenho. - Modelo Diana, S.Sebast. da Encruzilhada.

(*) - Método de Máxima Verossimilhança.

SÉRIE	SMA				SP1				SP2				SP3			
Nº \ T	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000	100	250	500	1.000
5	GU 0,24	GU 0,25	GU 0,25	GU 0,26	EXP* 0,24	EXP 0,25	EXP 0,26	GU 0,26	EXP* 0,21	EXP* 0,22	EXP 0,22	GU 0,22	EXP* 0,17	EXP* 0,17	EXP* 0,18	EXP* 0,19
10	GU 0,18	GU 0,19	GU 0,20	GU 0,20	EXP* 0,17	EXP* 0,18	EXP 0,18	EXP 0,19	EXP* 0,14	EXP* 0,15	GU 0,15	GU 0,15	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,13
15	GU 0,16	GU 0,17	GU 0,18	GU 0,19	EXP* 0,14	EXP* 0,15	EXP 0,16	EXP 0,17	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,14	EXP* 0,09	EXP* 0,09	EXP* 0,09	EXP* 0,10
20	GU 0,14	GU 0,15	GU 0,15	GU 0,16	EXP* 0,11	EXP* 0,12	EXP* 0,14	GU 0,14	EXP* 0,09	EXP* 0,10	EXP* 0,11	GU 0,11	EXP* 0,08	EXP* 0,08	EXP* 0,08	EXP* 0,08
25	GU 0,14	GU 0,14	GU 0,14	GU 0,15	EXP* 0,12	EXP* 0,13	EXP 0,14	GU 0,14	EXP* 0,09	EXP* 0,10	EXP* 0,11	GU 0,11	EXP* 0,07	EXP* 0,07	EXP* 0,07	EXP* 0,08
30	GU 0,12	GU 0,12	GU 0,13	GU 0,14	EXP* 0,09	EXP* 0,10	EXP* 0,11	EXP 0,11	EXP 0,08	EXP 0,09	EXP 0,10	GU 0,10	EXP* 0,07	EXP* 0,07	EXP* 0,07	EXP* 0,08
40	GU 0,10	GU 0,11	GU 0,11	GU 0,12	EXP* 0,08	EXP* 0,09	EXP* 0,10	EXP 0,11	EXP* 0,07	EXP* 0,08	EXP* 0,09	GU 0,09	EXP* 0,06	EXP* 0,06	EXP* 0,06	EXP* 0,07
50	GU 0,11	GU 0,11	GU 0,12	GU 0,13	EXP* 0,08	EXP* 0,09	EXP 0,10	EXP 0,11	EXP* 0,06	EXP* 0,06	EXP* 0,08	GU 0,09	EXP* 0,05	EXP* 0,05	EXP* 0,05	EXP* 0,06
60	GU 0,08	GU 0,09	GU 0,09	GU 0,09	EXP* 0,07	EXP* 0,08	EXP 0,09	EXP 0,10	EXP* 0,06	EXP 0,06	EXP 0,08	GU 0,09	EXP 0,05	EXP* 0,05	EXP* 0,05	EXP* 0,07

Tabela 4.19 - Erros Médios Absolutos das Distribuições de Melhor Desempenho. - Modelo Diana, Colatina

(*) - Método de Máxima Verossimilhança.

tuação se inverteu, tendo o método de máximos anuais apresentado resultados bem superiores aos dos métodos de séries parciais. Entretanto, como pode ser visto da tabela 4.18, os dois métodos apresentaram sub-estimativas dos valores de população, sendo que nos métodos de séries parciais os valores estimados foram da ordem de 70% dos de população. Para os métodos de máximos anuais, esta tendência foi um pouco menor, o que explica, em parte, a diferença dos desempenhos.

Foi tentado o ajuste de uma distribuição Pareto às séries parciais, mas os erros médios absolutos foram enormes, sendo os quantis de população invariavelmente superestimados. De qualquer forma, de nada adiantaria procurar uma distribuição que fornecesse bons resultados para este caso particular, onde a população é conhecida, a menos que fosse possível identificá-la a partir de características amostrais. Nota-se que este cenário foi o que apresentou máximos anuais com maiores desvio padrão, assimetria e curtose: entretanto, em amostras pequenas esta distinção nem sempre pode ser feita.

Enfim, embora os métodos de séries parciais tenham apresentado comportamentos melhores que os métodos de máximos anuais em 4 dos 5 cenários estudados, não é possível uma conclusão muito generalizada. O que pode ser considerado um resultado positivo destes experimentos é que o método de séries parciais pode ser útil para a avaliação de vazões com grandes tempos de recorrência e deve, por isso, ser considerado uma ferramenta de uso mais frequente pelos hidrólogos.

Para os 4 cenários do modelo DIANA, os experimentos acima foram repetidos com séries parciais obtidas pelo critério dos máximos locais (critério b item 2.3.2). Os resultados obtidos mostraram diferenças muito pequenas, com relação aos mostrados neste item, não alterando as conclusões.

4.5 - Testes de Adequação de Ajuste

Já foi mencionado o fato de que os testes de adequação de ajuste, freqüentemente utilizados como medida para a escolha da distribuição a ser utilizada na análise de freqüência de extremos, são capazes de quantificar apenas a capacidade de cada modelo de reproduzir as freqüências das vazões registradas no histórico. A princípio, nada se pode afirmar a respeito das acuidades destes modelos na previsão de vazões superiores (ou inferiores) às registradas.

Aproveitando a grande massa de dados oferecida pelos modelos de geração de vazões diárias aqui utilizados, foi feita uma experiência com o objetivo de se testar a validade da aplicação dos testes de adequação de ajuste quando se trata de escolher uma distribuição de probabilidades para aplicação em previsão de eventos extremos. A experiência pode ser descrita da seguinte forma:

1. Dividem-se as séries sintéticas de extremos disponíveis em sub-amostras de N elementos.
2. A cada sub-amostra são ajustadas as diferentes distribuições de probabilidades a serem estudadas.

3. A adequação do ajuste obtido para cada distribuição é medida pela estatística de Kolmogorov-Smirnov. Define-se a distribuição i que apresentou o melhor ajuste.
4. A partir de cada distribuição ajustada são estimadas as vazões com os tempos de recorrência desejados e, para cada tempo de recorrência, define-se a distribuição j que apresentou o menor erro $|\hat{q}(T) - q(T)|$.
5. Incrementa-se na posição (i,j) de uma matriz de resultados, "R", uma unidade. Passa-se à sub-amostra seguinte.

O valor do elemento $R(i,j)$, após a repetição deste processo para todas as sub-amostras, indica o número de vezes que a distribuição j fez a melhor estimativa da vazão extrema, tendo a distribuição i apresentado o melhor ajuste à amostra. Caso a matriz "R" apresente os elementos da diagonal principal superiores aos demais, o teste de Kolmogorov-Smirnov pode ser aplicado com sucesso na escolha da distribuição mais adequada para a estimação de vazões extremas.

Esta metodologia foi aplicada às séries sintéticas de máximos anuais geradas pelo modelo XADREZ. Foram tomadas sub-amostras de 5, 10, 15, 20, 25, 30, 40 e 50 elementos, às quais foram ajustadas, pelo método dos momentos, 4 distribuições de probabilidades de dois parâmetros: log-normal, exponencial, Gama e Gumbel. A tabela 4.20 mostra uma matriz "R" típica obtida para sub-amostras de 20 anos e tempo de recorrência de 10000 anos. Pode-se verificar que esta matriz não é diagonal-dominante, ou seja: o fato de que uma dada distribuição ajustou-se melhor a uma amostra

não implica, necessariamente, que esta distribuição seja a mais adequada para a estimação de vazões com tempos de recorrência elevados. Além disso, é possível notar que a distribuição marginal dos melhores ajustes não é parecida com a distribuição marginal das melhores estimativas de quantis. Deve-se enfatizar que este resultado foi sistemático, para tempos de recorrência de 100 a 10000 anos e para amostras de 5 a 50 elementos.

Para uma melhor documentação deste experimento, são apresentadas nas tabelas 4.21 a 4.28 todas as matrizes "R" obtidas. Para cada uma destas matrizes foi feita uma análise, como a exemplificada anteriormente com a tabela 4.20.

		MELHOR ESTIMATIVA DE QUANTIS				
		LN	EXP	GAM	GV	Σ
MELHOR AJUSTE	LN	19	31	57	41	148
	EXP	14	24	62	23	123
	GAM	18	40	31	29	118
	GV	6	54	23	28	111
	Σ	57	149	173	121	500

Tabela 4.20 - Matriz R obtida a partir de amostras de 20 elementos e tempo de recorrência 10000 anos.

O mesmo experimento foi realizado utilizando as séries parciais, apenas para considerar uma população diferente, tendo conduzido a resultados semelhantes aos aqui apresentados.

N = 5		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
	LN	45	168	136	34	28	184	134	37	26	183	134	40	22	163	144	54	383
	EXP	73	293	167	82	46	282	195	92	39	282	203	91	35	240	236	104	615
	GAM	73	318	131	58	49	338	132	61	48	343	126	63	42	313	139	86	580
	GU	38	272	83	29	32	276	74	40	29	276	72	45	24	262	77	59	422
	Σ	229	1.051	517	203	155	1.080	535	230	142	1.084	535	239	123	978	596	303	2.000

Tabela 4.21 - Matrizes "R" Obtidas para Amostras de 5 Anos:

OBS.: { 1 - melhor ajuste
2 - melhor estimativa de quantis

N = 10		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
	LN	27	112	93	22	23	112	93	26	21	111	90	32	21	91	99	43	254
	EXP	74	120	100	45	50	112	122	55	47	110	124	58	38	80	158	63	339
	GAM	32	90	54	18	26	95	45	28	24	96	42	32	15	88	49	42	194
	GU	20	150	23	20	20	150	18	25	21	149	15	28	27	128	23	35	213
	Σ	153	472	270	105	119	469	278	134	113	466	271	150	101	387	329	183	1.000

Tabela 4.22 - Matrizes "R" Obtidas para Amostras de 10 Anos:

OBS.: { 1 - melhor ajuste
2 - melhor estimativa de quantis

N = 15		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
LN		25	89	60	26	24	87	60	29	22	84	60	34	16	63	72	49	200
EXP		46	63	48	15	22	52	67	31	26	50	63	33	23	36	80	33	172
GAM		28	66	36	19	22	67	32	28	23	66	29	31	19	52	34	44	149
GU		9	94	20	22	9	91	19	26	9	90	19	27	22	65	21	37	145
Σ		108	312	164	82	77	297	178	114	80	290	171	125	80	216	207	163	666

Tabela 4.23 - Matrizes "R" Obtidas para Amostras de 15 Anos:

OBS.: { 1 - melhor ajuste
2 - melhor estimativa de quantis

N = 20		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
LN		19	55	47	27	21	49	45	33	24	49	45	30	19	31	57	41	148
EXP		32	39	34	18	19	38	47	19	18	36	47	22	14	24	62	23	123
GAM		26	50	29	13	25	45	25	23	24	46	24	24	18	40	31	29	118
GU		9	73	22	7	16	64	22	9	14	63	22	12	6	54	23	28	111
Σ		86	217	132	65	81	196	139	84	80	194	138	88	57	149	173	121	500

Tabela 4.24 - Matrizes "R" Obtidas para Amostras de 20 anos

OBS.: { 1 - melhor ajuste
2 - melhor estimativa de quantis

N = 25		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
	LN	20	52	40	18	25	45	34	26	23	46	32	29	17	28	43	42	130
	EXP	20	23	27	10	17	19	29	15	16	19	30	15	16	11	40	13	80
	GAM	12	45	14	17	13	41	11	23	12	41	11	24	12	32	17	27	88
	GU	9	55	28	10	11	49	24	18	10	48	24	20	17	33	29	23	102
	Σ	61	175	109	55	66	154	98	82	61	154	97	88	62	104	129	105	400

Tabela 4.25 - Matrizes "R" Obtidas para Amostras de 25 Anos:

OBS.: { 1 - melhor ajuste
2 - melhor estimativa de quantis

N = 30		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
	LN	17	30	39	20	16	26	37	27	16	26	37	27	11	12	51	32	106
	EXP	9	13	14	8	5	12	17	10	4	11	18	11	12	5	18	9	44
	GAM	13	46	13	11	14	44	11	14	15	43	11	14	15	29	13	26	83
	GU	9	48	24	19	14	44	18	24	15	43	14	28	16	32	26	26	100
	Σ	48	137	90	58	49	126	83	75	50	123	80	80	54	78	108	93	333

Tabela 4.26 - Matrizes "R" Obtidas para Amostras de 30 Anos:

OBS.: { 1 - melhor ajuste
2 - melhor estimativa de quantis

N = 40		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
LN		10	26	24	25	16	18	23	28	17	17	23	28	18	8	29	30	85
EXP		4	7	10	6	1	2	12	12	2	2	13	10	2	0	14	11	27
GAM		11	29	9	11	14	27	6	13	13	26	5	16	12	15	10	23	60
GU		12	34	20	12	13	29	18	18	12	28	17	21	18	12	24	24	78
Σ		37	96	63	54	44	76	59	71	44	73	58	75	50	35	77	88	250

Tabela 4.27 - Matrizes "R" Obtidas para Amostras de 40 Anos:

OBS.: $\begin{cases} 1 - \text{melhor ajuste} \\ 2 - \text{melhor estimativa de quantis} \end{cases}$

N = 50		T = 100				T = 500				T = 1.000				T = 10.000				Σ
1 \ 2		LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	LN	EXP	GAM	GU	
LN		25	22	23	17	21	19	21	26	20	19	20	28	14	10	35	28	87
EXP		3	2	5	4	3	1	6	4	2	1	7	4	2	1	9	2	14
GAM		8	17	3	9	10	13	0	14	12	12	0	13	7	10	4	16	37
GU		13	28	11	10	11	22	9	20	13	20	8	21	15	10	13	24	62
Σ		49	69	42	40	45	55	36	64	47	52	35	66	38	31	61	70	200

Tabela 4.28 - Matrizes "R" Obtidas para Amostras de 50 Anos:

OBS.: $\begin{cases} 1 - \text{melhor ajuste} \\ 2 - \text{melhor estimativa de quantis} \end{cases}$

CAPÍTULO V

CONCLUSÕES E RECOMENDAÇÕES

5.1 - Conclusões

As primeiras investigações realizadas tiveram o objetivo de caracterizar estatisticamente as séries parciais e discutir a aplicabilidade de algumas hipóteses freqüentemente assumidas nos modelos matemáticos destas séries. Estas investigações tomaram por base as séries de vazões médias diárias de 29 postos fluviométricos da bacia do rio Doce, para as quais foram definidas 4 séries de vazões extremas: a série de máximos anuais (SMA) e as séries parciais contendo em média 1, 2 e 3 picos por ano (SP1, SP2 e SP3). Os principais resultados obtidos são listados a seguir:

- A principal hipótese assumida pelos modelos de séries parciais é a da aleatoriedade das mesmas. Foram avaliados dois critérios de obtenção das séries parciais a partir dos registros de vazões médias diárias. Pelo primeiro critério, considera-se uma cheia como um conjunto de vazões que ultrapassem certo nível de truncamento previamente definido durante um intervalo contínuo de tempo, sendo a magnitude da cheia a máxima vazão ocorrida neste período. O segundo critério considera qualquer máximo local como um pico de cheia. Estes dois critérios revelaram-se equivalentes quanto à sua capacidade de fornecer séries parciais com pequeno grau de não aleatoriedade, pelo menos para séries com até três picos por ano, em média. Para taxas maiores de picos, o segundo critério começava a apresentar séries mais dependentes. O critério de má

ximos excessos (primeiro critério) foi considerado melhor, principalmente por ser conceitualmente mais representativo.

- A variável aleatória "número de picos a cada ano" apresentou um comportamento não compatível com a distribuição Poisson, o que contraria uma hipótese assumida em inúmeros modelos descritos na literatura. Entretanto, para a estimação de vazões com tempo de recorrência mais elevados, esta hipótese não é vital, pois, neste caso, é suficiente que se conheça o valor esperado do número de picos a cada ano.
- A variável aleatória "magnitude dos picos" apresentou estatísticas regionais razoavelmente próximas das da distribuição exponencial, confirmando uma hipótese clássica. As relações regionais assimetria/curtose das séries parciais da bacia do rio Doce são muito próximas desta relação teórica para as distribuições gama, sendo que para a taxa de um pico por ano esta relação é praticamente idêntica à da distribuição exponencial.
- A hipótese de que os picos tenham uma distribuição com cauda exponencial não pode ser rejeitada a nível regional pelo teste Gnedenko (Gnedenko et al., 1969). Entretanto, localmente esta hipótese foi rejeitada em poucos casos, de forma compatível com o nível de significância adotado.
- Os coeficientes regionais de variação, assimetria e curtose apresentaram valores crescentes à medida que foram tomadas séries parciais mais extensas (menor nível de truncamento). A faixa de variação destes coeficientes foi a seguinte, para λ

variando de um a três picos por ano:

- coeficiente de variação: 0,32 a 0,41
- coeficiente de assimetria: 2,01 a 2,72
- coeficiente de curtose: 8,80 a 13,52.

Embora o coeficiente de variação regional das séries de máximos anuais fosse superior aos das séries parciais, estas últimas apresentaram maiores assimetrias e curtoses que aquelas.

Estes primeiros pontos citados permitem concluir pela validade de se aplicar os métodos de séries parciais na análise de frequência de cheias, uma vez que as principais hipóteses das mesmas parecem razoáveis. Não foi investigada a hipótese de que os picos sejam provenientes de uma mesma população, o que é intuitivamente falso, devido à sazonalidade. Pode-se admitir, no entanto, que os picos sejam provenientes de uma distribuição mista (Bardsley), o que permite um tratamento analítico bem mais simples.

Quanto ao aproveitamento de informações regionais para aplicação dos modelos de séries parciais, foram tentadas algumas abordagens, considerando a bacia do rio Doce como uma região hidrologicamente homogênea. As principais conclusões obtidas são descritas a seguir:

- As regressões lineares associando a média e o nível de truncamento das séries parciais às áreas das bacias hidrográficas apresentaram excelentes resultados, melhores que a regressão obtida para as médias dos máximos anuais. Observou-se que o coeficiente de determinação das regressões melhora à medida que se tomam séries parciais mais extensas, ou seja: níveis

de truncamento mais baixos.

- A tentativa de obterem-se distribuições regionais para as séries parciais da bacia do rio Doce não obteve êxito. Foram tentadas três distribuições regionais: Wakeby, valores extremos generalizada e exponencial. Em todos os casos, as distribuições regionais obtidas apresentaram um mau ajuste às séries parciais de vários postos. Este resultado poderia ser tomado como indício de não homogeneidade da bacia do rio Doce. Entretanto, como para as séries de máximos anuais a distribuição Wakeby regional apresentou resultados muito bons, é possível que esta aparente não homogeneidade, verificada nas séries parciais, seja decorrente de uma abordagem equivocada do tratamento regional para as séries parciais.
- A origem deste equívoco poderia ser o tratamento indistinto dado às séries parciais, provenientes de bacias com áreas de drenagem muito diferentes. Entretanto, mesmo dividindo-se os postos em classes de áreas de drenagem semelhantes, não se conseguiu obter distribuições regionais mais satisfatórias.

Foram criados cenários constituídos por um grande número de séries sintéticas de vazões diárias, a partir dos quais pode ser feita uma comparação do desempenho dos métodos de máximos anuais e dos métodos de séries parciais, no tocante à determinação de vazões com tempos de recorrência elevados. Os resultados revelaram que o desempenho dos dois métodos depende da população, ou seja, as conclusões listadas a seguir devem ser entendidas como conclusões específicas para cada caso. Isso não impede que se tente generalizar alguns resultados, mas exige, no entanto,

que as generalizações sejam feitas com cautela.

O primeiro cenário, constituído de 100.000 séries de vazões diárias, permitiu resultados mais precisos da aplicação da metodologia de avaliação dos desempenhos dos métodos. Para este cenário, os principais resultados foram:

- Os métodos de máximos anuais e de séries parciais tiveram desempenhos semelhantes na estimação de vazões com pequeno tempo de recorrência (até 50 anos).
- Para a estimativa de vazões com tempos de recorrência superiores a 100 anos, os métodos de séries parciais tiveram desempenhos superiores aos métodos de máximos anuais, principalmente no caso de amostras pequenas (menos de 30 anos).
- Quanto maior a taxa anual de picos das séries parciais, menores os erros médios absolutos verificados nas estimativas de quantis.
- As distribuições que apresentaram os menores erros médios absolutos nas estimativas de quantis foram a exponencial, nos métodos de séries parciais, e a Gumbel, nos métodos de máximos anuais. Este resultado é animador, pois, teoricamente, era esperado. De maneira geral, exceto a distribuição normal, que apresentava sub-estimativas dos quantis, todas as distribuições de dois parâmetros ofereceram resultados bons.

Os quatro cenários restantes, constituídos de 10.000 séries de vazões diárias, representativos de registros flu

viométricos da bacia do rio Doce, forneceram estimativas menos precisas dos erros médios absolutos. Entretanto, permitiram as conclusões expostas a seguir:

- Para três dos quatro cenários, os métodos de séries parciais apresentaram desempenho superior ao dos métodos de máximos anuais, confirmando os resultados do primeiro experimento.
- Para o cenário representativo do registro do posto de São Sebastião da Encruzilhada, no Rio Manhuaçu, o método de séries parciais conduziu a erros médios absolutos muito grandes, bem superiores aos erros dos métodos de máximos anuais.
- Este resultado pode ser atribuído à grande assimetria da população, o que conduziu à subestimação dos quantis, tanto a partir das séries parciais, quanto a partir das séries de máximos anuais. Para este cenário, em vários casos (amostras maiores) as distribuições de três parâmetros permitiram erros médios absolutos mais baixos que as de dois parâmetros.
- As séries parciais obtidas pelos dois critérios estudados (máximos excessos e máximos locais) se prestaram à análise de extremos de forma equivalente. A escolha entre os dois critérios de obtenção pode ser considerada uma questão apenas formal. Sugere-se, no entanto, o método de máximos excessos por apresentar significado conceitual mais claro.

A escolha do modelo probabilístico a ser utilizado para a análise de frequência de cheias, seja por séries parciais, seja por máximos anuais, deve ser feita a partir de evidências a-

mostrais. Os testes de adequação de ajuste são comumente usados como base para a escolha deste ou daquele modelo para análise de frequência de cheias. Entretanto, esses testes avaliam apenas a eficiência dos modelos em reproduzir as frequências dos eventos amostrais, nada garantindo acerca do comportamento destes modelos na estimação de vazões extremas.

- Foi aplicada uma metodologia de enumeração para verificar a eficiência dos testes de adequação de ajuste no tocante à seleção de uma distribuição para a estimação de vazões com tempos de recorrência elevados. Concluiu-se que estes métodos podem levar a resultados equivocados, pois a distribuição que melhor se ajusta a uma amostra não é, necessariamente, a que fará melhor as extrapolações.

Finalmente, não foi possível concluir, tomando por base as experiências realizadas, qual a série de extremos que conduz a resultados mais confiáveis na determinação de vazões com tempos de recorrência elevados. Sendo esta resposta dependente da população, não se pode, a priori, fazer uma opção, sendo mais prudente utilizar mais de um tipo de série na análise direta de frequência de extremos.

5.2 - Recomendações

Tendo em vista os resultados obtidos neste estudo, podem ser feitas as seguintes recomendações:

- Recomenda-se a utilização dos métodos de séries parciais como método auxiliar para a estimação de vazões com tempos de

recorrência elevados, considerando que, dado o grau de incerteza que envolve a extrapolação requerida neste caso, é sempre válido dispor de mais um resultado. Nos casos de se dispor de amostras pequenas, o método de séries parciais pode oferecer resultados mais confiáveis que os métodos de máximos anuais.

- As metodologias de regionalização, recomendáveis para a estimação de vazões com grande tempo de recorrência, devem ser aplicadas com cuidado no caso de séries parciais. O bom resultado de regionalização das séries de máximos anuais não garante o mesmo para as séries parciais.
- É recomendável o uso de distribuições de dois parâmetros para a estimação de vazões com tempos de recorrência muito elevados. Quanto à escolha da distribuição de dois parâmetros a ser utilizada, os testes de cauda podem fornecer resultados úteis, quando aplicados num sentido regional.
- Não se devem utilizar os testes de adequação de ajuste como medida para a escolha da distribuição de probabilidades para grandes extrapolações. Estes testes, eficientes para avaliar o desempenho dos modelos probabilísticos em reproduzirem frequências amostrais, podem conduzir a resultados equivocados se aplicados à escolha de modelos para a estimação de vazões extremas.

Algumas limitações deste estudo devem ser lembradas, com os objetivos de fixar o alcance das conclusões e recomendações apresentadas e de sugerir algumas extensões das metodolo-

gias propostas.

- As aplicações das metodologias de séries parciais foram feitas apenas para séries da bacia do rio Doce. Seria interessante a aplicação destas metodologias a outras bacias hidrográficas do Brasil.
- Não foi possível identificar a origem do mau resultado obtido na tentativa de se chegar a uma distribuição regional para as séries parciais. Seria recomendável verificar se este mau resultado se repete para outras bacias.
- As conclusões quanto aos desempenhos dos métodos de séries parciais e de máximos anuais se prendem à população que forneceu as séries. Neste trabalho, foram utilizadas séries estocásticas de vazões diárias para a criação destas populações. Seria interessante a aplicação da metodologia de avaliação proposta a cenários de vazões diárias obtidas a partir de modelos conceituais de chuva-deflúvio.
- Finalmente, recomenda-se a aplicação da metodologia aqui proposta, para o estudo de séries parciais que considerem os volumes, picos e durações das cheias.

BIBLIOGRAFIA

- ASHKAR, F. e J. ROUSSELLE - "Some remarks on the truncation used in partial flood series models", Water Resources Research, vol. 19, nº 2, 1983.
- ASHKAR, F. e J. ROUSSELLE - "The effect of certain restrictions imposed on the interarrival times of flood events on the Poisson Distribution used for modeling flood counts", Water Resources Research, vol. 19, nº 2, 1983.
- BARBOSA, P. J. - "Modelagem estocástica de vazões diárias", Tese de Mestrado submetida à COPPE/UFRJ, 1981.
- BARDSLEY, W. E. e F. J. MANLY - "Note on the distribution of flood maxima for random sample size", Journal of Hidrology, vol. 42, pág. 187-193, 1979.
- BARDSLEY, W. E. - "Note on the application of classical extreme value theory to flood data", Journal of Hydrology, vol. 42, pág. 195-198, 1979.
- BENJAMIN, J. R. e C.A. CORNELL - "Probability, statistics and decision for civil engineers", McGraw-Hill Book Company, New York, 1970.
- BORGMAN, L.E. - "Risk criteria", Journal of Waterways and Harbor Division, ASCE, vol. 89, pág. -35, WW3, 1963.

- BRADLEY, J. F. - "Distribution-free statistical tests", Prentice Hall, New Jersey, 1968.
- CUNNANE, C. - "A note on the Poisson assumption in partial duration series models", Water Resources Research, vol. 15, nº 2, 1979.
- DAMÁZIO, J. M. e J. KELMAN - "Análise não paramétrica de séries de vazões extremas", Relatório Técnico CEPEL 154/83, 1983.
- DAMÁZIO, J.M., J.C. MOREIRA, J.P. DA COSTA e J. KELMAN - "Seleção de método para estimação de vazões com tempos de retorno elevados", Relatório Técnico CEPEL 156/83, 1983.
- EPSTEIN, B. - "A modified extreme value problem", Ann. Math.Stat., vol. 20, pág. 99-103, 1949.
- FISHER, R. A. e L.H.C. TIPPETT - "Limiting forms of the frequency distribution of the largest or smallest member of a sample", Proc. Cambridge Philos. Soc., vol. 24, pág. 180-190, 1928.
- GNEDENKO, B., Y. K. BELYAYEV e A. D. SOLOVYEV - "Mathematical methods of reliability theory", Probability and Mathematical Statistics Series, Academic Press, New York, 1969.
- GOMIDE, F.L.S. e E. KAVISKI - "Poder de testes de adequação de ajuste: uma investigação experimental", IV Simpósio Brasileiro de Hidrologia e Recursos Hídricos, Fortaleza, 1981.

GUMBEL, E. J. - "Statistics of extremes", Columbia University Press, New York, 1958.

GUPTA, V. K., L. DUCKSTEIN e R. W. PEEBLES - "On the joint distribution of the largest flood and its time of occurrence", Water Resources Research, vol. 12, nº 2, April 1976.

HOUGHTON, J. C. - "Birth of a parent: the Wakeby Distribution for modeling flood flows", Water Resources Research, vol. 14, nº 6, 1978.

KARR, A. F. - "Two extreme value processes arising in Hydrology" Journal of Applied Probability, vol. 13, pág. 190-194, 1976.

KAVVAS, M. L. - "Stochastic trigger model for flood peaks: 1. Development of the model, 2. Application of the model to the flood peaks of Goksu-Karahacili", Water Resources Research, vol. 5, nº 4, August 1982.

KELMAN, J., J. M. DAMÁZIO e J. P. DA COSTA - "Geração de séries sintéticas de vazões diárias: Metodologia do Modelo DIANA, Relatório Técnico CEPEL 120/83, 1983.

KIRBY, W., - "On the random occurrence of major floods", Water Resources Research, Vol. 5, nº 4, August 1969.

- KOTTEGODA, N.T. - "Tail behavior of river flow probability densities", Proc. Int. Symp. on the Risk and Reliability in Water Resources, Ontario, 1979.
- KOTTEGODA, N. T. - "Stochastic water resources technology", The Macmillan Press Ltd, London, 1980.
- KUCZERA, G. - "Robust flood frequency models", Water Resources Research, vol. 18, n° 2, 1982.
- KYTE, G. W. - "Frequency and risk analyses in Hydrology", Water Resources Publications, Fort Collins, Co., 1977.
- LANDWEHR, J. M., N. C. MATALAS e J. R. WALLIS - "Estimation of parameters and quantiles of Wakeby Distribution. Part I: Known Lower Bounds", Water Resources Research, vol. 15, n° 6, 1979.
- LANDWEHR, J.M., N.C. MATALAS e J. R. WALLIS - "Estimation of parameters and quantiles of Wakeby Distribution. Part II: Unknown Lower Bounds", Water Resources Research, vol. 15, n° 6, 1979.
- LANDWEHR, J. M., N. C. MATALAS e J. R. WALLIS - "Quantile estimation with more or less floodlike distributions", Water Resources Research, vol. 16, n° 3, 1980.
- MOOD, A.M., F. A. GRAYBILL e D. C. BOES - "Introduction to the theory of statistics", third edition, McGraw-Hill, New York, 1974.

NATIONAL ENVIRONMENT RESEARCH COUNCIL-NERC - "Flood studies report", London, 1975.

NORTH, M. - "Time-dependent stochastic model of floods", Journal of Hydraulics Division, ASCE, vol. 106, HY5, pág. 649-665, May 1980.

PINHEIRO, M.C. - "Análise de Extremos e Regionalização de Quantis". Tese de Mestrado submetida à COPPE/UFRJ em novembro de 1982.

SHANE, R. e W. LYNN - "Mathematical model for flood risk evaluation", Journal of Hydraulics Division, ASCE, vol. 90, HY6, pág. 1-20, June 1964.

TODOROVIC, P. - "On some problems involving random number of random variables", Ann. Math. Statistic, vol. 41(3), pág., 1059-1063, 1970.

TODOROVIC, P. e ZELENHASIC, A. - "A stochastic model for flood analysis", Water Resources Research, vol. 6(6), pág. 1641-1658, 1970.

TODOROVIC, P. e ROUSSELLE, J. - "Some problems of flood analysis" Water Resources Research, vol. 7(5), pág. 1144-1150, 1971.

TODOROVIC, P. e D. A. WOOLHISER - "On the time when the extreme flood occurs", Water Resources Research, vol. 8(6), pág. 1433-1438, 1972.

- TODOROVIC, P. e H. W. SHEN - "Some remarks on the statistical theory of extremes values, in stochastic approaches to Water Resources", vol. 2, edited by H. W. Shen, University of Colorado Press, Fort Collins, Co., 1976.
- TODOROVIC, P. - "Stochastic models of floods", Water Resources Research, vol. 14(2), pág. 345-356, 1978.
- WALLIS, J. R. - "Risk and uncertainties in the evaluation of flood events for the design of hydrologic structures, Seminar on extreme hydrological events: floods and droughts", Erice, Italia, março de 1980.
- WATER RESOURCES CONUNCIL - "Guidelines for determining flood flow frequency", Bull. 17, Hydrol. Comm. Washington, D.C., 1976.
- YEVJEVCH, V. e V. TAESOMBUT - "Information on flood peaks in daily flow series", Proc. Int. Symp. on Risk and Reliability in Water Resources, Ontario, Canada, 1979.
- ZELENHASIC, E. - "Theoretical probability distributions for flood peaks", Colorado State University Hydrology Papers, nº 42, Colorado State University, Fort Collins, 1970.

APÊNDICE AESTIMAÇÃO DE QUANTIS

Seja X uma variável aleatória da qual se tem n observações. Defina-se

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$s^2 = \frac{n}{n-1} \left(\sum_{i=1}^n x_i^2 - \bar{x}^2 \right)$$

$$g = \left(\frac{n^2}{(n-1)(n-2)s^3} \left(\frac{\sum_{i=1}^n x_i^3}{n} - \frac{3\bar{x} \sum_{i=1}^n x_i^2}{n} + \frac{3\bar{x} \sum_{i=1}^n x_i}{n} - \bar{x}^3 \right) \right)$$

como estimadores da média, variância e assimetria respectivamente. De modo análogo, define-se \bar{x}' , s'^2 e g' como média, variância e assimetria dos logaritmos da variável aleatória substituindo-se x_i por $\log x_i$ nas fórmulas acima.

A seguir apresenta-se o cálculo do quantil de projeto q_T associado ao período de recorrência T (e portanto com probabilidade de ser excedido $p = 1 - \frac{1}{T}$) para as diversas distribuições de probabilidades utilizadas neste estudo.

1. Distribuição Normal

$$q_T = \bar{x} + z.s$$

onde $z = \Phi^{-1}(p)$ dado por

$$z = t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} ; \quad t = \sqrt{\ln \frac{1}{p}}$$

para $0 < p \leq 0,5$

Se $p > 0,5$ use $(1-p)$ e troque o sinal de z

$$c_0 = 2,515517$$

$$d_1 = 1,432788$$

$$c_1 = 0,802853$$

$$d_2 = 0,189269$$

$$c_2 = 0,010328$$

$$d_3 = 0,001308$$

2. Distribuição log normal

$$q_T = \exp (\bar{x}' + z s')$$

$$z = \Phi^{-1}(p)$$

3. Distribuição log normal a três parâmetros

$$q_T = \exp (\mu_n + z \sigma_n) + c_n$$

onde

$$\mu_n = \ln \frac{s}{c_1} - 0,5 \ln (c_2^2 + 1)$$

$$\sigma_n = \sqrt{\ln (c_2^2 + 1)}$$

$$c_n = \bar{x} \left(1 - \frac{c_1}{c_2} \right)$$

$$c_1 = \frac{s}{\bar{x}}$$

$$c_2 = \frac{1 - w^{2/3}}{w^{1/3}}$$

$$w = \frac{-|g| + \sqrt{(g^2 + 4)}}{2}$$

$$z = \Phi^{-1}(p)$$

Se $g < 0,007$, considerou-se que a distribuição é normal.

4. Distribuição exponencial (método dos momentos)

$$q_T = x_0 - \beta \cdot \ln(1/T)$$

onde

$$x_0 = \bar{x} - s$$

$$\beta = s$$

5. Distribuição exponencial (método da máxima verossimilhança)

$$q_T = x_0 - \beta \ln(1/T)$$

$$x_0 = \min_i x_i - \frac{\beta}{n}$$

$$\beta = \frac{n}{n-1} (\bar{x} - \min_i x_i)$$

6. Distribuição Pearson III

$$q_T = \alpha \beta \left(1 - \frac{1}{9\beta} + z \sqrt{\frac{1}{9\beta}} \right)^3 + c$$

onde

$$\beta = \frac{4}{g^2}$$

$$\alpha = \frac{s}{\sqrt{\beta}} \operatorname{sign}(g)$$

$$c = \bar{x} - \alpha \beta$$

$$z = \Phi^{-1}(p) \operatorname{sign}(g)$$

Se $|g| < 0,007$, considerou-se distribuição normal

7. Distribuição log-Pearson III

$$q_T = \exp(q_T')$$

onde q_T' é obtido como no item 6 a partir de \bar{x}' , s' e g'

8. Distribuição Gumbel

$$q_T = \mu - \alpha \left(\ln(-\ln(p)) \right)$$

onde

$$\alpha = 0,78.s$$

$$\mu = \bar{x} - 0,577.\alpha$$

9. Distribuição generalizada de valores extremos

$$q_T = \mu + \frac{\alpha}{k} \left(1 - (-\ln(p))^k \right)$$

onde

$$\mu = \bar{x} + \frac{\alpha}{k} \left(\Gamma(k+1) - 1 \right)$$

$$\alpha = \frac{|k \cdot s|}{\left(\Gamma(2k+1) - \left(\Gamma(k+1) \right)^2 \right)^{1/2}}$$

$\Gamma(\cdot)$: função gama

k: obtido por interpolação linear em função de g

g	k
- 0,3581	0,4
- 0,2166	0,35
- 0,0688	0,30
+ 0,0866	0,25
0,2558	0,20
0,4362	0,15
0,623	0,10
0,9115	0,05
1,532	- 0,05
1,903	- 0,10
2,532	- 0,15
3,535	- 0,20
5,605	- 0,25

Se $|k| \leq 0,001$, adotou-se a distribuição Gumbel

10. Distribuição Gama

$$q_T = \alpha \beta \left(1 - \frac{1}{9\beta} + z \sqrt{\frac{1}{9\beta}} \right)^3, \quad q_T \geq 0$$

onde

$$\alpha = s^2 / \bar{x}$$

$$\beta = \bar{x}^2 / s^2$$

$$z = \Phi^{-1}(p)$$

11. Distribuição Wakeby

$$q_T = m + a \left(1 - \left(\frac{1}{T} \right)^b \right) - c \left(1 - \left(\frac{1}{T} \right)^{-d} \right)$$

O método utilizado para estimação dos parâmetros requer o uso dos momentos ponderados pela probabilidade, M_k dados por:

$$M_k = \frac{1}{n} \sum_{i=1}^n x_i (1-F_i)^k \quad k = 0, 1, 2, 3, 4$$

$$F_i = \frac{i - 0,35}{n}, \quad x_1 < x_2 < \dots < x_n$$

Uma vez obtido M_k usou-se o algoritmo descrito por Landwehr et al.(1979) como "caminho" $[V_m, m=0]$ para estimar os parâmetros m, a, b, c e d .

APÊNDICE BDESCRIÇÃO DO MODELO XADREZB.1 - Introdução

Apresenta-se uma breve descrição do modelo XADREZ, desenvolvido por Jerson Kelman (comunicação pessoal). Este método foi utilizado para a geração de séries sintéticas de vazões médias diárias.

O modelo considera que a vazão afluente a um posto fluviométrico num certo dia t é dada por:

$$Q(t) = Q(t-1) + X(t) ; t = 2,3, \dots, h \quad (B.1)$$

Dada uma série histórica $q(1), q(2), \dots, q(365)$, pode-se obter, a partir da equação B.1, uma série de incrementos $x(2), x(3), \dots, x(365)$, de forma que:

$$x(t) = q(t) - q(t-1) ; t = 2,3, \dots, 365 \quad (B.2)$$

Partindo da equação B.1, pode-se construir um modelo gerador de hidrógrafas sintéticas, procedendo da seguinte forma:

- i) define-se um valor inicial $q(1)$
- ii) sorteia-se aleatoriamente e com reposição 364 valores do conjunto $x(t); t=1,2, \dots, 364$ definido pela equação B.2.

Este modelo simples permite a obtenção de $(364)^{364}$ hidrógrafas, todas elas igualmente prováveis.

Na prática, são disponíveis n hidrógrafas históricas

$$q_i(t) ; t = 1, 2, \dots, 365$$

$$i = 1, 2, \dots, n;$$

sendo possível a definição de um conjunto de incrementos com $364n$ valores, a partir da equação B.2. Além disso, pode-se dispor de um conjunto de valores iniciais, $q_i(1) ; i = 1, 2, \dots, n$.

O número de hidrógrafas que podem ser obtidas pelo algoritmo anterior - apenas com a modificação do primeiro passo, pois o valor inicial passa a ser sorteado também - cresce para $n(364n)^{364}$. Todas as hidrógrafas são ainda igualmente prováveis.

A aceitação deste procedimento é equivalente à aceitação da hipótese de que a série

$$x_i(t) ; t = 1, 2, \dots, 365;$$

$$i = 1, 2, \dots, n$$

é resultante de 364 sorteios independentes de uma mesma população suficientemente representada pela distribuição empírica dos $364n$ elementos.

Sabe-se, entretanto, que as séries de vazões diárias apresentam fortes componentes de persistência e sazonalidade

Este modelo simples descrito pode ser aperfeiçoado para levar em conta estas duas componentes.

CONSIDERAÇÃO DA SAZONALIDADE

Particiona-se o conjunto $x(t); t = 2, 3, \dots, 365;$
 $i = 1, 2, \dots, n$ em $\frac{364}{m}$ sub-conjuntos de m elementos:

$$x_i(t); t = 1, 2, \dots, m; i = 1, 2, \dots, n,$$

$$x_i(t); t = m + 1, m + 2, \dots, 2m; i = 1, 2, \dots, n,$$

$$\vdots$$

$$x_i(t); t = 364 - m + 1, 364 - m + 2, \dots, 364; i = 1, 2, \dots, n$$

onde m pode ser, por exemplo, um mês.

Neste caso, as hidrôgrafas sintéticas seriam obtidas através de $\frac{364}{m}$ sequências de m sorteios aleatórios, feitos em cada sub-conjunto definido acima. O número de hidrôgrafas sintéticas possíveis diminui agora para $n(\frac{364}{m} n)^{\frac{364}{m}}$, sendo todas elas ainda equiprováveis.

CONSIDERAÇÃO DA PERSISTÊNCIA

O efeito da persistência pode ser incorporado ao modelo por meio da divisão de cada um dos m sub-conjuntos definidos anteriormente em dois grupos, contendo um, os elementos (incrementos) positivos, e o outro, os elementos (incrementos) negativos. Uma cadeia de Markov é associada ao modelo através dos parâ

metros:

$$\pi = P [X(t) \geq 0 \mid X(t-1) \geq 0] \quad (B.3)$$

$$\phi = P [X(t) < 0 \mid X(t-1) < 0] \quad (B.4)$$

Com este procedimento, o número de diferentes hidrógrafas sintéticas a serem fornecidas pelo modelo não se altera, mas algumas delas passam a ser mais prováveis do que outras. A probabilidade do sorteio de uma certa hidrógrafa é uma função complexa de π e ϕ .

Pode-se adotar ainda uma classificação adicional para os grupos de sorteio,

$$x_i(t); t = 1, 2, \dots, 365; i = 1, 2, \dots, n,$$

de acordo com a vazão total anterior $q(t-1)$. Esta classificação é útil para levar em conta o fato de que certos incrementos diários de vazões são mais prováveis durante uma enchente que durante uma estiagem, por exemplo.

B.2 - O Modelo Xadrez

A aplicação do modelo XADREZ exige que se classifique o conjunto de incrementos diários de vazões do histórico,

$$x_i(t); t = 2, 3, \dots, 365; i = 1, 2, \dots, n,$$

de acordo com três critérios, definidos a seguir:

Critério A: $a = 1$ se $x(t) \geq 0$
 $a = 2$ se $x(t) < 0$

Critério B: $b = j$ se $q_{j-1} < q(t-1) < q_j$

Critério C: $c = i$ se $\tau_{i-1} < t < \tau_i$

Para cada par (b,c) , é definida uma matriz de transição da cadeia de Markov, através dos parâmetros $\pi(b,c)$ e $\phi(b,c)$. Estes parâmetros são obtidos por enumeração das transições dos valores de a .

A geração de hidrógrafas sintéticas é feita de acordo com o seguinte algoritmo:

- i) $t = 1$: - sorteia-se $q(1)$ da distribuição empírica das vazões iniciais; $a=1$
- ii) $t = t+1$
- iii) define-se o valor de b de acordo com $q(t-1)$ e o de c de acordo com t ;
- iv) sorteia-se u da distribuição uniforme entre zero e um:
- v) se $a = 2$, vá até vii
- vi) se $u > \pi(b,c)$, então $a = 2$
vã até viii
- vii) se $u > \phi(b,c)$, então $a = 1$
- viii) sorteia-se $x(t)$ da distribuição empírica da classe (a,b,c)
- ix) $q(t) = q(t-1) + x(t)$

x) se t não é 365, vá até ii.

B.3 - Calibração do Modelo

A calibração do modelo consiste na determinação dos vetores $\underline{q} = (q_0, q_1, \dots, q_r)$ e $\underline{\tau} = (\tau_0, \tau_1, \dots, \tau_m)$, de forma tal que se possam obter séries sintéticas com propriedades semelhantes às das séries históricas.

Este processo de calibração está concebido como um jogo de tentativa e erro. Inicialmente, é feita uma análise preliminar dos dados e, por inspeção, definem-se os componentes do vetor \underline{q} que devem dividir as vazões em altas ou baixas. Da mesma forma, definem-se as componentes do vetor $\underline{\tau}$, que devem dividir o ano em estações diferentes.

Definidos estes valores iniciais \underline{q} e $\underline{\tau}$, manipulam-se os dados das vazões, de modo a definir as matrizes de transição e os vetores de incrementos para cada par (b,c) .

Esta primeira calibragem é testada gerando-se algumas séries e comparando suas propriedades (algumas estatísticas) com as do histórico.

É importante frisar que, quanto maior a dimensão dos vetores \underline{q} e $\underline{\tau}$, mais parecidos com o histórico serão as séries geradas. Entretanto, o número de hidrógrafas possíveis diminui bastante, e a probabilidade de que se sorteie uma hidrógrafa diferente das hidrógrafas históricas (com um pico maior, por exemplo) cai bastante. Deve-se procurar, então, trabalhar com as menores di-

mensões possíveis para os vetores \underline{q} e $\underline{\tau}$.